

99925

STIC-Biotech/ChemLib

From: Chan, Christina  
Sent: Monday, July 28, 2003 6:04 PM  
To: Sullivan, Daniel; STIC-Biotech/ChemLib  
Subject: RE: RUSH Sequence Search for 09834291

Please rush. Thanks Chris

Chris Chan  
TC 1600 New Hire Training Coordinator and SPE 1644  
308-3973  
CM-1, 9B19

-----Original Message-----

From: Sullivan, Daniel  
Sent: Monday, July 28, 2003 1:45 PM  
To: Chan, Christina  
Subject: RUSH Sequence Search for 09834291  
Importance: High

Hi Chris-

Could you please approve the following search for an after final amended case.

Please search for the following in the commercial, interference and issued patent databases:

A nucleic acid sequence comprising SEQ ID NO: 1, 2, 3, 4, 6, 15 or 24;

A nucleic acid sequence consisting of SEQ ID NO: 10, 12 or 14

Thank you.

Daniel M. Sullivan  
Examiner AU 1636  
Room: 12D12  
Mail Box: 11E12  
Tel: 703-305-4448

Searcher: Jan  
Phone: 4657  
Location: \_\_\_\_\_  
Date Picked Up: 7/29/03  
Date Completed: 8/3/03  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: 20  
Online time: 15

TYPE OF SEARCH: ☒  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \*  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_



Sun Aug 3 09:03:33 2003

us-09-834-291-1.rge

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:37:27 ; Search time 12552.1 Seconds  
(without alignment)  
10468.541 Million cell updates/sec

Title: US-09-834-291-1

Perfect score: 1 tgcagactcgcagatg.....ttgcagatggtcaatcaag 3212

Sequence: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Scoring table: 2888711 seqs, 20454813366 residues

Searched: 5777422

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenBank1: \*  
1: gb\_ba: \*  
2: gb\_hcg: \*  
3: gb\_in: \*  
4: gb\_ov: \*  
5: gb\_ov: \*  
6: gb\_ov: \*  
7: gb\_ov: \*  
8: gb\_ov: \*  
9: gb\_ov: \*  
10: gb\_ov: \*  
11: gb\_ov: \*  
12: gb\_ov: \*  
13: gb\_ov: \*  
14: gb\_ov: \*  
15: gb\_ov: \*  
16: gb\_ov: \*  
17: gb\_ov: \*  
18: gb\_ov: \*  
19: gb\_ov: \*  
20: gb\_ov: \*  
21: gb\_ov: \*  
22: gb\_ov: \*  
23: gb\_ov: \*  
24: gb\_ov: \*  
25: gb\_ov: \*  
26: gb\_ov: \*  
27: gb\_ov: \*  
28: gb\_ov: \*  
29: gb\_ov: \*  
30: gb\_ov: \*  
31: gb\_ov: \*  
32: gb\_ov: \*  
33: gb\_ov: \*  
34: gb\_ov: \*  
35: gb\_ov: \*  
36: gb\_ov: \*  
37: gb\_ov: \*  
38: gb\_ov: \*  
39: gb\_ov: \*  
40: gb\_ov: \*  
41: gb\_ov: \*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match length	DB ID	Description
1	3212	100.0	3212 6	AX026089
2	3193	99.4	45121 6	AX026089
3	3185	99.2	187313 6	AX026089
4	2346.6	73.1	2827 6	AX026089
5	2211.8	68.9	2344 6	AX026089
6	2143.4	68.7	2165 6	AX026089
7	1899.6	59.1	1877 9	AX026089
8	1850.8	57.6	1877 9	AX026089
9	1536.8	47.8	1608 6	AX026089
10	961	29.9	1608 6	AX026089
11	957.4	22.2	720 6	AX026089
12	713	22.2	702 9	AX026089
13	698.8	21.8	702 9	AX026089
14	697.2	21.7	702 9	AX026089
15	695.6	21.7	702 9	AX026089
16	431.2	13.4	436 9	AX026089
17	429.6	13.4	436 9	AX026089
18	342	10.6	266 6	AX026089
19	262.8	8.2	266 6	AX026089
20	257	8.0	398 11	AX026089
21	257	8.0	398 11	AX026089
22	252	7.8	251 6	AX026089
23	252	7.8	251 6	AX026089
24	252	7.8	251 6	AX026089
25	252	7.8	251 6	AX026089
26	252	7.8	251 6	AX026089
27	226	7.0	2471 6	AX026089
28	226	7.0	2471 6	AX026089
29	226	7.0	2471 6	AX026089
30	226	7.0	2471 6	AX026089
31	226	7.0	2471 6	AX026089
32	226	7.0	2471 6	AX026089
33	226	7.0	2471 6	AX026089
34	226	7.0	2471 6	AX026089
35	208.6	6.5	1728 4	AX026089
36	78.6	2.5	1728 4	AX026089
37	78.6	2.5	1728 4	AX026089
38	78.6	2.5	1728 4	AX026089
39	78.6	2.5	1728 4	AX026089
40	78.6	2.5	1728 4	AX026089
41	78.6	2.5	1728 4	AX026089
42	78.6	2.5	1728 4	AX026089
43	78.6	2.5	1728 4	AX026089
44	78.6	2.5	1728 4	AX026089
45	78.6	2.5	1728 4	AX026089

ALIGNMENTS

RESULT 1  
LOCUS AX026089 3212 bp  
DEFINITION Sequence 1 from Patent DE19847779.  
ACCESSION AX026089  
VERSION AX026089.1 GI:10187520  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Mueller-Schilling, M., Kramer, P. and Oren, M.  
TITLE Novel receptor dna useful for identifying apoptosis-modulating  
substances potentially useful for cancer chemotherapy

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: DE 19847779-C 1 03-FEB-2000;  
FEATURES DEUTSCHES KREBSFORSCH (DE)  
Source Location/Qualifiers  
1..3212

BASE COUNT /organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN 778 a 784 c 809 g 841 t

Query Match 100.0%; Score 3212; DB 6; Length 3212;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TGAGGACTCTGAGAAATATGCTGTAATAATAATAATACTTTAGATGCCCAACTGT 60  
1 TGAGGACTCTGAGAAATATGCTGTAATAATAATAATACTTTAGATGCCCAACTGT 60  
61 TTTCCCGAAGAACACGACATTCATTTAGATTTCAATGATTTCTTCAAGATTTCCA 120  
61 TTTCCCGAAGAACACGACATTCATTTAGATTTCAATGATTTCTTCAAGATTTCCA 120  
121 AAGGCAAGAAAGTTTGGGGAACATATATATATATATATATATATATATATATAT 120  
121 AAGGCAAGAAAGTTTGGGGAACATATATATATATATATATATATATATATATAT 120  
181 TAAGGACCTGAGAAATTTGGATTAAGAAAGTTTCAATTAAGTAACCCAGAAATTT 180  
181 TAAGGACCTGAGAAATTTGGATTAAGAAAGTTTCAATTAAGTAACCCAGAAATTT 180  
241 CTAGATTTATTTGACATGTAACATATATATATATATATATATATATATATATAT 240  
241 CTAGATTTATTTGACATGTAACATATATATATATATATATATATATATATATAT 240  
301 GAACTTGAGATTAATTAAGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 300  
301 GAACTTGAGATTAATTAAGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 300  
361 AAGAGCAGAGACCTTGGGGAACAAATATATATATATATATATATATATATATAT 360  
361 AAGAGCAGAGACCTTGGGGAACAAATATATATATATATATATATATATATATAT 360  
421 ACTAACCACTTTTCCGATGCTTAAAGCTTTTGGCTACATTTTATTTTATTTAAG 420  
421 ACTAACCACTTTTCCGATGCTTAAAGCTTTTGGCTACATTTTATTTTATTTAAG 420  
481 TAAGTTATATATACATCTATCTACATCTATCTATCTATCTATCTATCTATCTAT 480  
481 TAAGTTATATATACATCTATCTACATCTATCTATCTATCTATCTATCTATCTAT 480  
541 CACATATATGATGCTGCTTATATATATATATATATATATATATATATATATAT 540  
541 CACATATATGATGCTGCTTATATATATATATATATATATATATATATATATAT 540  
601 CTTTCCCTTTTCT 600  
601 CTTTCCCTTTTCT 600  
661 CTTCCCTGACAGCCCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660  
661 CTTCCCTGACAGCCCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660  
721 ATTTCGACAGCCCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720  
721 ATTTCGACAGCCCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720  
781 TACCATCT 780  
781 TACCATCT 780  
841 GGTGGGACAGCCGAGGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840  
841 GGTGGGACAGCCGAGGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840  
841 GGTGGGACAGCCGAGGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840  
841 GGTGGGACAGCCGAGGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840

901 ACAAGCCATCTACACCTTACAGACTGTTGTTATGTCAGTACATGATGCAAAACACAG 960  
901 ACAAGCCATCTACACCTTACAGACTGTTGTTATGTCAGTACATGATGCAAAACACAG 960  
961 GTGATGGAAGCCCTCAGAGAGGTAACTTAACCTTATGATTTAGGAGCCCAAGCTTCAG 960  
961 GTGATGGAAGCCCTCAGAGAGGTAACTTAACCTTATGATTTAGGAGCCCAAGCTTCAG 960  
1021 AAGAAATGTCACTGAGAGAGAGCTTGAAGATGAACAGTGGCTTAACAAAGGTTAT 1020  
1021 AAGAAATGTCACTGAGAGAGAGCTTGAAGATGAACAGTGGCTTAACAAAGGTTAT 1020  
1081 TAATGTTATTAATGCTGTTAACTTAATTTGGGAGAGAGAGGTTGCAAGTGAAGT 1080  
1081 TAATGTTATTAATGCTGTTAACTTAATTTGGGAGAGAGAGGTTGCAAGTGAAGT 1080  
1141 CAGAGCTTGTGAGACATGCAAGAGAAATCTGAACCTTTAGTGTCCAGTGTGAAC 1140  
1141 CAGAGCTTGTGAGACATGCAAGAGAAATCTGAACCTTTAGTGTCCAGTGTGAAC 1140  
1201 TGATCCAAATTCAGGTTCAATATGATGATGATGATGATGATGATGATGATGAT 1200  
1201 TGATCCAAATTCAGGTTCAATATGATGATGATGATGATGATGATGATGATGAT 1200  
1261 ATGCTAACTACCTTAAGCTATCTACCGTTGCAAGCAATGATGATGATGATGATG 1260  
1261 ATGCTAACTACCTTAAGCTATCTACCGTTGCAAGCAATGATGATGATGATGATG 1260  
1321 CACAGAGCAGAGAAAGATTAACAATTTTATTTTAAAGAAATTTGGCCAGAAATATG 1320  
1321 CACAGAGCAGAGAAAGATTAACAATTTTATTTTAAAGAAATTTGGCCAGAAATATG 1320  
1381 AGTAAGAAAGACAGAAAGTATTTGATGATGATGATGATGATGATGATGATGATG 1380  
1381 AGTAAGAAAGACAGAAAGTATTTGATGATGATGATGATGATGATGATGATGATG 1380  
1441 CTTAAGTTAGAGCTTTTCTCTTGAAGAAATTAAGAGGCTTCCCTTTTTC 1440  
1441 CTTAAGTTAGAGCTTTTCTCTTGAAGAAATTAAGAGGCTTCCCTTTTTC 1440  
1501 AGAGCTTATAGGCGCAACATCTGATCTTTTCAATATGTTAATCTGATCTCAAGAAAG 1500  
1501 AGAGCTTATAGGCGCAACATCTGATCTTTTCAATATGTTAATCTGATCTCAAGAAAG 1500  
1561 TCTGTAGGCTCTCATGTTGAGAGCAACATGACAGAGCCCAAGTCCCGCAAGT 1560  
1561 TCTGTAGGCTCTCATGTTGAGAGCAACATGACAGAGCCCAAGTCCCGCAAGT 1560  
1621 CTTTCTGAGAGTCTCCAGCAATTAAGCAAGGCTCTGTAACCCAGAGAGAGCTCTG 1620  
1621 CTTTCTGAGAGTCTCCAGCAATTAAGCAAGGCTCTGTAACCCAGAGAGAGCTCTG 1620  
1681 CTTGAGCTCATTTCTCTTCAAGACCTTCCCAATTTCCAGGTGAATCAAGAGAG 1680  
1681 CTTGAGCTCATTTCTCTTCAAGACCTTCCCAATTTCCAGGTGAATCAAGAGAG 1680  
1741 CTTTGAAGAGGAGAGAGAGGAGGCTCTGAGGTCTTCACTCAAGGTGAATCAAGAGAG 1740  
1741 CTTTGAAGAGGAGAGAGAGGAGGCTCTGAGGTCTTCACTCAAGGTGAATCAAGAGAG 1740  
1801 ACTGAG 1800  
1801 ACTGAG 1800  
1861 CACCTGATCTTCT 1860  
1861 CACCTGATCTTCT 1860  
1921 CTTCCCAACCCGAGGCTTCCAGAGAGGCTTCTTCCATCTCTGACACAGAGGCT 1920  
1921 CTTCCCAACCCGAGGCTTCCAGAGAGGCTTCTTCCATCTCTGACACAGAGGCT 1920  
1980 CTTCCCAACCCGAGGCTTCCAGAGAGGCTTCTTCCATCTCTGACACAGAGGCT 1980  
1980 CTTCCCAACCCGAGGCTTCCAGAGAGGCTTCTTCCATCTCTGACACAGAGGCT 1980



QY	1981	TTTTGCTGAGCTGCTCTCTGATCTCGGCAAGATGACACAGGGTTCMAAAGCCTTC	2040
Db	1981	TTTTGCTGAGCTGCTCTCTGATCTCGGCAAGATGACACAGGGTTCMAAAGCCTTC	2040
QY	2041	TGGGAGTGAAGGGAAGGGGTTTACAGATGACTTGGCTGGAAGCTCAGAGGGCGGGCACTGG	2100
Db	2041	TGGGAGTGAAGGGAAGGGGTTTACAGATGACTTGGCTGGAAGCTCAGAGGGCGGGCACTGG	2100
QY	2041	TGGGAGTGAAGGGAAGGGGTTTACAGATGACTTGGCTGGAAGCTCAGAGGGCGGGCACTGG	2100
Db	2041	TGGGAGTGAAGGGAAGGGGTTTACAGATGACTTGGCTGGAAGCTCAGAGGGCGGGCACTGG	2100
QY	2101	CACGGAAACACCCCTGAGGCCAGGCCCTGGGCGCCAGAGCGAGCTGCTCTCCCGG	2160
Db	2101	CACGGAAACACCCCTGAGGCCAGGCCCTGGGCGCCAGAGCGAGCTGCTCTCCCGG	2160
QY	2161	GGTTGGTGAACCCGCTCAGTACGAGTTGGAGAACTCTTCACTTCGAGATTTGCTCA	2220
Db	2161	GGTTGGTGAACCCGCTCAGTACGAGTTGGAGAACTCTTCACTTCGAGATTTGCTCA	2220
QY	2221	ACAAACATGCTGGGCAATCGAACCTCTCACTCTGTGTATCCCTCTCGTCCGGGTGG	2280
Db	2221	ACAAACATGCTGGGCAATCGAACCTCTCACTCTGTGTATCCCTCTCGTCCGGGTGG	2280
QY	2281	AGGCTTACCCCGCTTAACTCCCGGGGATAGGCAAGTGGGCGGGGCGGAGCGGTGG	2340
Db	2281	AGGCTTACCCCGCTTAACTCCCGGGGATAGGCAAGTGGGCGGGGCGGAGCGGTGG	2340
QY	2341	GGATTGGCGGCGACAGGCGGCAACCTGGAGACCGCGGCTGTCCGGAGAGGG	2400
Db	2341	GGATTGGCGGCGACAGGCGGCAACCTGGAGACCGCGGCTGTCCGGAGAGGG	2400
QY	2341	GGATTGGCGGCGACAGGCGGCAACCTGGAGACCGCGGCTGTCCGGAGAGGG	2400
Db	2341	GGATTGGCGGCGACAGGCGGCAACCTGGAGACCGCGGCTGTCCGGAGAGGG	2400
QY	2401	TTGAGAGCTGGGCTCCGGGGGCTGTAGAACCTTCCCTCAGGGCCGGGTGCTCAGAAACA	2460
Db	2401	TTGAGAGCTGGGCTCCGGGGGCTGTAGAACCTTCCCTCAGGGCCGGGTGCTCAGAAACA	2460
QY	2461	TGAGAGACTTGTCTTTCTTGGGCTTGATGCGAAGTGTGATCCGCTGGGCGAGCGGG	2520
Db	2461	TGAGAGACTTGTCTTTCTTGGGCTTGATGCGAAGTGTGATCCGCTGGGCGAGCGGG	2520
QY	2521	CAGCTCCGGCGCTCTCCGAGACACATGCGCTTCAACGTGAGGTGGCGTGGAGGGCGGA	2580
Db	2521	CAGCTCCGGCGCTCTCCGAGACACATGCGCTTCAACGTGAGGTGGCGTGGAGGGCGGA	2580
QY	2581	CAGAAATTGAAGCGGAAGCTCTGGAACAAGCTTGAAGGTGGCGTGGAGGGCGGA	2640
Db	2581	CAGAAATTGAAGCGGAAGCTCTGGAACAAGCTTGAAGGTGGCGTGGAGGGCGGA	2640
QY	2641	GAGAGAGCGGAACCTCTGGAACAAGCTTGAAGGTGGCGTGGAGGGCGGA	2700
Db	2641	GAGAGAGCGGAACCTCTGGAACAAGCTTGAAGGTGGCGTGGAGGGCGGA	2700
QY	2701	GGGTGGGTGAGTGGCGCGCGCGCGCGCGGGGGGGGAGAGGCTTCAAGCTTCAAGACA	2760
Db	2701	GGGTGGGTGAGTGGCGCGCGCGCGCGCGGGGGGGGAGAGGCTTCAAGCTTCAAGACA	2760
QY	2761	CATATTGCTCATTTCTGGAAGTCTGAGAGCTTGAAGAAATTAAGTCAGACCGAAGAGTG	2820
Db	2761	CATATTGCTCATTTCTGGAAGTCTGAGAGCTTGAAGAAATTAAGTCAGACCGAAGAGTG	2820
QY	2821	GTTAAGCGGAGGGCTCGGAAGACGCACTTTTCTTTCGAAAAAGTTATATGGGG	2880
Db	2821	GTTAAGCGGAGGGCTCGGAAGACGCACTTTTCTTTCGAAAAAGTTATATGGGG	2880
QY	2881	CTGATGAGCTTCTGAGAGCTTGAACGCTTTTATTTTATTTGTCACACAGAAAAAGAACTG	2940
Db	2881	CTGATGAGCTTCTGAGAGCTTGAACGCTTTTATTTTATTTGTCACACAGAAAAAGAACTG	2940
QY	2941	CGTGTCTCCCTCCGGGAATTTCTCTTTAAGACTGTAAGTCCCTCCGAGAGGTTTC	3000
Db	2941	CGTGTCTCCCTCCGGGAATTTCTCTTTAAGACTGTAAGTCCCTCCGAGAGGTTTC	3000
QY	3001	ATTGTTGTTTGTCTTCTGAGCTTCTCTTCTTCTTGGGCTTCTTACCTTCACTCC	3060
Db	3001	ATTGTTGTTTGTCTTCTGAGCTTCTCTTCTTCTTGGGCTTCTTACCTTCACTCC	3060
QY	3061	CATGATGATTTCTGCTTGGTCTCTGCGGGGTGGTGTACTGTTCCACCGCAGACA	3120
Db	3061	CATGATGATTTCTGCTTGGTCTCTGCGGGGTGGTGTACTGTTCCACCGCAGACA	3120

[illegible]



Where differences are found these are indicated together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emi, EMBL; Swi, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; information on the WORMPEP



Db	121	AAGGCAAGAAAGAAATTCAGCAATTC	240
Qy	181	TAAAGGCCCCCTGAGAAAGTTTGGATTGAAGAAGTTTCAATTAAAGTAAACCAAGAAATTT	240
Db	181	TAAAGGCCCCCTGAGAAAGTTTGGATTGAAGAAGTTTCAATTAAAGTAAACCAAGAAATTT	240
Qy	241	CTAAGATTAATTTGACCATATGAAACATATGTCTCCCAACAAAGACATATTTCTATCTCTT	300
Db	241	CTAAGATTAATTTGACCATATGAAACATATGTCTCCCAACAAAGACATATTTCTATCTCTT	300
Qy	301	GAACCTTAGAGATTAATTAGAAGTACAGGAGTAAAGGGTAAAGGGGAGTATGGCATAG	360
Db	301	GAACCTTAGAGATTAATTAGAAGTACAGGAGTAAAGGGTAAAGGGGAGTATGGCATAG	360
Qy	361	AAGAGCAGAGACCTTGGAGCAGAAATATCTAAGTTAATTCCTGACTCTGCTATTATTA	420
Db	361	AAGAGCAGAGACCTTGGAGCAGAAATATCTAAGTTAATTCCTGACTCTGCTATTATTA	420
Qy	421	ACTAACCAATCTTGGCAATGTGCTTAAGCTTTTGGCTACATTTTAAATTTTAAAG	480
Db	421	ACTAACCAATCTTGGCAATGTGCTTAAGCTTTTGGCTACATTTTAAATTTTAAAG	480
Qy	481	TAAATTAATTAATCACTCATATCACTGGGCTATATATATATATATATATATATATAT	540
Db	481	TAAATTAATTAATCACTCATATCACTGGGCTATATATATATATATATATATATATATAT	540
Qy	541	CACATATGATGATGTGCTTATTAATCAACCAAGAGATATCTGATTTGTCAATTTGT	600
Db	541	CACATATGATGATGTGCTTATTAATCACTCAAGAGATATCTGATTTGTCAATTTGT	600
Qy	601	CGTTTCCCTTTTCTCTCTTCCCTCTTCATTCCTCTCCCTAACCTCTCTCTT	660
Db	601	CGTTTCCCTTTTCTCTCTTCCCTCTTCATTCCTCTCCCTAACCTCTCTCTT	660
Qy	661	CTTCCCTCAACCCCTTCTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	720
Db	661	CTTCCCTCAACCCCTTCTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	720
Qy	721	ATTATGATTAATTTAGATTTCAAAAATTTGACAGATATATACAGAAATGCCCTA	780
Db	721	ATTATGATTAATTTAGATTTCAAAAATTTGACAGATATATACAGAAATGCCCTA	780
Qy	781	TACCATCTCTCTTATCCCACTCTTTTGTGTCTATTAATGATCTCAGAGTGTGACAA	840
Db	781	TACCATCTCTCTTATCCCACTCTTTTGTGTCTATTAATGATCTCAGAGTGTGACAA	840
Qy	841	GAGTGGACGCCAGGCT	900
Db	841	GAGTGGACGCCAGGCT	900
Qy	901	ACAACCCATCAACCTCTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	960
Db	901	ACAACCCATCAACCTCTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	960
Qy	961	GTATGTAAGAGCCCTCAAGAGGTATCTTAACCTTAAGTGAAGGCCCAACAGGCTCCA	1020
Db	961	GTATGTAAGAGCCCTCAAGAGGTATCTTAACCTTAAGTGAAGGCCCAACAGGCTCCA	1020
Qy	1021	GAGAAATATGTCATCTGAGAGAGAGCTGAGATGACAGTGTGCTTAAGCAAGTGAAGT	1080
Db	1021	GAGAAATATGTCATCTGAGAGAGAGCTGAGATGACAGTGTGCTTAAGCAAGTGAAGT	1080
Qy	1081	TAAATGTTAATTAATGAGTGAATCTAATTTGGAGAGAGAGAGAGAGAGAGAGAGT	1140
Db	1081	TAAATGTTAATTAATGAGTGAATCTAATTTGGAGAGAGAGAGAGAGAGAGAGAGT	1140
Qy	1141	GCAGAGCTTGTGAGAGATGCAAGAGATCTGAACCTTAAGTGTCTCAAGCTGGAA	1200
Db	1141	GCAGAGCTTGTGAGAGATGCAAGAGATCTGAACCTTAAGTGTCTCAAGCTGGAA	1200
Qy	1201	CTGCATCCAAATTCAGTTCAATTAATGATGTCTATTCAAACATATCTTCTGTAAAT	1260
Db	1201	CTGCATCCAAATTCAGTTCAATTAATGATGTCTATTCAAACATATCTTCTGTAAAT	1260





QY CACATATGAGTGTGCTGCTTATTAATTCAGACCTCAGAGATACAGTATTTGTCAATGTG 600  
541 |||||  
Db CACATATGAGTGTGCTGCTTATTAATTCAGACCTCAGAGATACAGTATTTGTCAATGTG 717  
658 |||||  
QY CCTTTCCCTTTTTTCTCTCTCCCTCCCTGCATTCCTCTCCCTTACCTCTCTCTTC 660  
601 |||||  
Db CCTTTCCCTTTTTTCTCTCTCCCTCCCTGCATTCCTCTCCCTTACCTCTCTCTTC 717  
718 CTTTTCCCTTTTTTCTCTCTCCCTCCCTGCATTCCTCTCCCTTACCTCTCTCTTC 720  
QY CTTTTCCCTTTTTTCTCTCTCCCTCCCTGCATTCCTCTCCCTTACCTCTCTCTTC 837  
661 CTTTTCCCTCACCCCTTTTCTCTCTCTTTTAACTTTTATTAAAGCAATTTTC 837  
Db CTTTTCCCTCACCCCTTTTCTCTCTCTTTTAACTTTTATTAAAGCAATTTTC 780  
718 CTTTTCCCTCACCCCTTTTCTCTCTCTTTTAACTTTTATTAAAGCAATTTTC 837  
QY ATTTTGAATAGTTTNGAATTCAAAATTTCAGAGATAATACAGAGATGCCATA 897  
721 ATTTTGAATAGTTTNGAATTCAAAATTTCAGAGATAATACAGAGATGCCATA 897  
Db ATTTTGAATAGTTTNGAATTCAAAATTTCAGAGATAATACAGAGATGCCATA 840  
838 ATTTTGAATAGTTTNGAATTCAAAATTTCAGAGATAATACAGAGATGCCATA 840  
QY TACCATTCCTTATCCCACTTCCTTTTGTGCTATTTAGATGCTCAGAGTGTGCACAA 957  
781 TACCATTCCTTATCCCACTTCCTTTTGTGCTATTTAGATGCTCAGAGTGTGCACAA 957  
||| TACCATTCCTTATCCCACTTCCTTTTGTGCTATTTAGATGCTCAGAGTGTGCACAA 960

Accession	Sequence	Length
QY	902	
QY	903	
QY	904	
QY	905	
QY	906	
QY	907	
QY	908	
QY	909	
QY	910	
QY	911	
QY	912	
QY	913	
QY	914	
QY	915	
QY	916	
QY	917	
QY	918	
QY	919	
QY	920	
QY	921	
QY	922	
QY	923	
QY	924	
QY	925	
QY	926	
QY	927	
QY	928	
QY	929	
QY	930	
QY	931	
QY	932	
QY	933	
QY	934	
QY	935	
QY	936	
QY	937	
QY	938	
QY	939	
QY	940	
QY	941	
QY	942	
QY	943	
QY	944	
QY	945	
QY	946	
QY	947	
QY	948	
QY	949	
QY	950	
QY	951	
QY	952	
QY	953	
QY	954	
QY	955	
QY	956	
QY	957	
QY	958	
QY	959	
QY	960	
QY	961	
QY	962	
QY	963	
QY	964	
QY	965	
QY	966	
QY	967	
QY	968	
QY	969	
QY	970	
QY	971	
QY	972	
QY	973	
QY	974	
QY	975	
QY	976	
QY	977	
QY	978	
QY	979	
QY	980	
QY	981	
QY	982	
QY	983	
QY	984	
QY	985	
QY	986	
QY	987	
QY	988	
QY	989	
QY	990	
QY	991	
QY	992	
QY	993	
QY	994	
QY	995	
QY	996	
QY	997	
QY	998	
QY	999	
QY	1000	
QY	1001	
QY	1002	
QY	1003	
QY	1004	
QY	1005	
QY	1006	
QY	1007	
QY	1008	
QY	1009	
QY	1010	
QY	1011	
QY	1012	
QY	1013	
QY	1014	
QY	1015	
QY	1016	
QY	1017	
QY	1018	
QY	1019	
QY	1020	
QY	1021	
QY	1022	
QY	1023	
QY	1024	
QY	1025	
QY	1026	
QY	1027	
QY	1028	
QY	1029	
QY	1030	
QY	1031	
QY	1032	
QY	1033	
QY	1034	
QY	1035	
QY	1036	
QY	1037	
QY	1038	
QY	1039	
QY		

QY	1080	TTAATGAGT	1199
Db	1198	TTAATGAGTATTAAATGGAATGGAGGAGCAAGGTTGCAAGAGT	1199
QY	1140	GCAGAGCTGTGGACGATGCCAAGAAATCTGAATCCTTTAGTGTGCCAGTCTGGA	1259
Db	1258	GCAGAGCTGTGGACGATGCCAAGAAATCTGAATCCTTTAGTGTGCCAGTCTGGA	1259
QY	1200	CTGCATCCAAATTCAGGTTCAATGATGTCATATCCAAACATACCTTCGTAAATTT	1377
Db	1318	CTGCATCCAAATTCAGGTTCAATGATGTCATATCCAAACATACCTTCGTAAATTT	1377
QY	1260	CATGCTAAATCTACTAAGATCTATCCGTTCCAAAGGATATGAGACTTTGAACAGTGT	1437
Db	1378	CATGCTAAATCTACTAAGATCTATCCGTTCCAAAGGATATGAGACTTTGAACAGTGT	1437
QY	1320	TCACAGAGCAGCAAGAAATTAAGAAATTTTTTTTAAAGAAATTTGGCAGAGAAATAT	1497
Db	1438	TCACAGAGCAGCAAGAAATTAAGAAATTTTTTTTAAAGAAATTTGGCAGAGAAATAT	1497
QY	1380	GAGTAAAGAGGACAGGAAATGATGTAATGTTAATATGAGCTGGGCTATGCGAATTTG	1557
Db	1498	GAGTAAAGAGGACAGGAAATGATGTAATGTTAATATGAGCTGGGCTATGCGAATTTG	1557
QY	1440	GCTTAAGTGTAAAGCTTTGTTTCCCTTGAGAAATATAAACTAAGAGGGCCCTCCCTTTT	1617
Db	1558	GCTTAAGTGTAAAGCTTTGTTTCCCTTGAGAAATATAAACTAAGAGGGCCCTCCCTTTT	1617
QY	1500	CAGAGCCTTATGGCGCAACATCTGTAATCTTTTCAATATGTTTAACTGTGCAATCCAGAAAC	1677
Db	1618	CAGAGCCTTATGGCGCAACATCTGTAATCTTTTCAATATGTTTAACTGTGCAATCCAGAAAC	1677
QY	1560	GTTCTGTAGCTCTCAATGTTGACAGCACAACATGGAACAGGCCAGTCAAAATGCCCGCAAG	1733
Db	1678	GTTCTGTAGCTCTCAATGTTGACAGCACAACATGGAACAGGCCAGTCAAAATGCCCGCAAG	1733

QY 1620 TCTTTCTCTGAGTACTCCAGCAATTGACCAAGGCTCTGTACCCAGCGAGACTCTGC 1679  
 Db 1738 TCTTCTCTGAGTACTCCAGCAATTGACCAAGGCTCTGTACCCAGCGAGACTCTGC 1679  
 QY 1680 GCTCTGAGCTCCCAATTCTCTCTTAAGACTCCCCCACTTCCAGGTGAATCAACAGCAAA 1739  
 Db 1798 GCTCTGAGCTCCCAATTCTCTCTTAAGACTCCCCCACTTCCAGGTGAATCAACAGCAAA 1739  
 QY 1740 GCGTTTGAAGAAGGGCAGAGGCGGCTCTCGAGAGCTCTCACTTGAAGGAGCATGCCAGC 1857  
 Db 1858 GCGTTTGAAGAAGGGCAGAGGCGGCTCTCGAGAGCTCTCACTTGAAGGAGCATGCCAGC 1799  
 QY 1800 CACTGCAGAAAGCGCCCGGGCAGAAATGCCCAATTGTGCAAGAAACCTGACTCTTCC 1917  
 Db 1918 CACTGCAGAAAGCGCCCGGGCAGAAATGCCCAATTGTGCAAGAAACCTGACTCTTCC 1859  
 QY 1860 TCACCTGACTCTTCCCTCCCTCCCTACCCGCGGCGAGCCAAATTGCTGAATCAATGAGC 1977  
 Db 1978 TCACCTGACTCTTCCCTCCCTCCCTACCCGCGGCGAGCCAAATTGCTGAATCAATGAGC 1919  
 QY 1920 CCTCCCCCAACCCGGGCGTTCCCAAGAGGCTTCTTCCACTCCCTGACACACCGGAGC 2037  
 Db 2038 CTTCCCCCAACCCGGGCGTTCCCAAGAGGCTTCTTCCACTCCCTGACACACCGGAGC 1979  
 QY 1980 TTTTGTGAGTCTGCTCTGTATCTGCGCAGAGTGCACACAGGTTCAAGACGGTT 2097  
 Db 2098 TTTTGTGAGTCTGCTCTGTATCTGCGCAGAGTGCACACAGGTTCAAGACGGTT 2039  
 QY 2040 CTGGGAGTGAAGGAAAGCGTTTACGAGTACTTGGCTGAGACTTCAGAGGGCGGCACTG 2157  
 Db 2158 CTGGGAGTGAAGGAAAGCGTTTACGAGTACTTGGCTGAGACTTCAGAGGGCGGCACTG 2099  
 QY 2100 GCACGGAAACACCTCGAGGCGAGCCGCTGCGCCAGCGGAGCTGCTCTTCTCCGCG 2159  
 Db 2218 GCACGGAAACACCTCGAGGCGAGCCGCTGCGCCAGCGGAGCTGCTCTTCTCCGCG 2277  
 QY 2160 GGGTTGATGACCCGCTCATGACGAGTTGGGGAAGCTCTTCACTTCGAGAGATTGCTC 2219  
 Db 2278 GGGTTGATGACCCGCTCATGACGAGTTGGGGAAGCTCTTCACTTCGAGAGATTGCTC 2219  
 QY 2220 AACAAAC 2226  
 Db 2338 AACAAAC 2344

	protein_bind	/organism="Homo sapiens"
	protein_bind	/mol_type="genomic DNA"
	protein_bind	/db_xref="taxon:9606"
	protein_bind	/feature="pfam"
	protein_bind	/feature_type="blood"
	protein_bind	79..87
	protein_bind	/bound_motety="NF-IL6"
	protein_bind	276..284
	protein_bind	/bound_motety="NF-IL6"
	protein_bind	481..489
	protein_bind	/bound_motety="NF-IL6"
	protein_bind	772..780
	protein_bind	/bound_motety="NF-IL6"
	protein_bind	934..940
	protein_bind	/bound_motety="AP-1"
	protein_bind	948..955
	protein_bind	/bound_motety="Ets"
	protein_bind	1288..1306
	protein_bind	/bound_motety="NF-IL6"
	protein_bind	1436..1813
	protein_bind	/number=1
	protein_bind	1436..1813
	protein_bind	1782..1813
	protein_bind	/codon_start=1
	protein_bind	/product="Pis antigen"
	protein_bind	/protein_id="BAIA20850.1"
	protein_bind	/db_xref="nr:4433150"
	protein_bind	/translation="MLGITLLPLV"
	protein_bind	1814..>2215
	protein_bind	/number=1
BASE COUNT	508 a	555 c
ORIGIN	555 g	547 t

Query Match	66.7%;	Score 2143.4;	DB 9;	Length 2165
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 2158; Conservative				

[illegible]



987 CCTTAACTAGATTGAGGCCC-AAACAGGCTCCAGAAAGAAATATGCACTGAGAGAAAC 1045  
541 CTTAACCTAGATTGAGGCCCCAAACAGGCTCCAGAAAGAAATATGCACTGAGAGAAAC 600  
1046 CTGAAGAGTGAACAGTGGCTTAAGCAAAAGGTTATTAATGTTATTAATGAGTTGAATC 1105  
601 CTGAAGAGTGAACAGTGGCTTAAGCAAAAGGTTATTAATGTTATTAATGAGTTGAATC 660  
1106 TAATTGGAGAGAGAGAGGTTGACAGTGAAGGTGACAGACTTGGTGAACATGCAAG 1165  
661 TAATTGGAGAGAGAGAGGTTGACAGTGAAGGTGACAGACTTGGTGAACATGCAAG 720  
1166 GAATCTGAACCTTATGATGATGTCAGTCTGGAACCTGCAATCCAAATTCAGTTCAAT 1225  
721 GAATCTGAACCTTATGATGATGTCAGTCTGGAACCTGCAATCCAAATTCAGTTCAAT 780  
1226 GATGTCATTATCCAAACATACCTCTGTAAATTCATGCTAAATCTAAGAGCTAATC 1285  
781 GATGTCATTATCCAAACATACCTCTGTAAATTCATGCTAAATCTAAGAGCTAATC 840  
1286 ACCGTTCCAAAGCAATAGTGAATTTGAACAGTGTACACAGAGCAGAAAGATTAAG 1345  
841 ACCGTTCCAAAGCAATAGTGAATTTGAACAGTGTACACAGAGCAGAAAGATTAAG 900  
1346 ATTTTATTTTAAAGAAATTTGGCAGAGAAATATAGTATACAGAGACAGAAATTAAG 1405  
901 ATTTTATTTTAAAGAAATTTGGCAGAGAAATATAGTATACAGAGACAGAAATTAAG 960  
1406 TGAATGTTATATATAGCTGGGCGTATGCAATTTGGCTTAAGTTGTTAGTTGTTTCT 1465  
961 TGAATGTTATATATAGCTGGGCGTATGCAATTTGGCTTAAGTTGTTAGTTGTTTCT 1020  
1466 CTTGGAATATTAATACTAAGGAGGCTCTCCCTTTTCAAGGCTTATAGGAGCAATCTGA 1525  
1021 CTTGGAATATTAATACTAAGGAGGCTCTCCCTTTTCAAGGCTTATAGGAGCAATCTGA 1080  
1526 CTTTTCATATGTTAATCTGTCAATTCAGAAAGCTCTGAGGCTCTCATGTTGAGGC 1585  
1081 CTTTTCATATGTTAATCTGTCAATTCAGAAAGCTCTGAGGCTCTCATGTTGAGGC 1140  
1586 ACAACATGACAGCCCAATGCAATGCCCCGCAAGCTTCTCTGAGTGAATCCAGCAAT 1645  
1141 ACAACATGACAGCCCAATGCAATGCCCCGCAAGCTTCTCTGAGTGAATCCAGCAAT 1200  
1646 AGCCAGAGCTCTGTACCCAGAGAGAGCTCTGAGGCTCTGAGCTCATTTCTTCAAG 1705  
1201 AGCCAGAGCTCTGTACCCAGAGAGAGCTCTGAGGCTCTGAGCTCATTTCTTCAAG 1260  
1706 CTTCCCAACTTCCAGAGTTGAATCAAGCAAGCTTTTAAAGAGGAGAGGCGGC 1765  
1261 CTTCCCAACTTCCAGAGTTGAATCAAGCAAGCTTTTAAAGAGGAGAGGCGGC 1320  
1766 TCTGAGGTCCTCACTGAAGTGAAGTGCAGCACTGAGAGAGCGCCCCGGAGCAAGA 1825  
1321 TCTGAGGTCCTCACTGAAGTGAAGTGCAGCACTGAGAGAGCGCCCCGGAGCAAGA 1380  
1826 ATGCCATTTGTCAGACGAACTCTGCTTCTTCAACCTGACCTTCCCTCCCTAC 1885  
1381 ATGCCATTTGTCAGACGAACTCTGCTTCTTCAACCTGACCTTCCCTCCCTAC 1440  
1886 CGCGGCGAGAGCCAGATTGCTGAATCAATGAGGCTTCCCAACCGGCGGCTTCCAGC 1945  
1441 CGCGGCGAGAGCCAGATTGCTGAATCAATGAGGCTTCCCAACCGGCGGCTTCCAGC 1500  
1946 GAGGCTCTTCTCCATCTCTCTGACACCGGGGCTTTTGATGCTCTCTCTGATCTCG 2005  
1501 GAGGCTCTTCTCCATCTCTCTGACACCGGGGCTTTTGATGCTCTCTCTGATCTCG 1560  
2006 CGCAAGATGACACAGAGTTCCTCAAGAGCTTCTGGAGATGAGAGAGAGGCTTTACG 2065  
1561 CGCAAGATGACACAGAGTTCCTCAAGAGCTTCTGGAGATGAGAGAGAGGCTTTACG 1620  
2066 ATGTGACTTGGCTGAGGCTCAAGGAGCGGCACTGACAGGAGACACCTCTGAGCGAGC 2125

	DB	1621	AGTCACTTGCGTGAACCTCAGGGGCGGCATCGGCACAACCAACCCTAGAAGCCACGCC	1680
	OY	2126	CTGCTCCCCAAGCGAGCTGCTCTTTCTCCGCGGTGGAGCACCGCTCAGTACGGA	2185
	DB	1681	CTGCTGCGCAAGCGAGCTGCTCTTTCTCCGCGGTGGAGCACCGCTCAGTACGGA	1740
	OY	2186	GTTGGGGAAGCTCTTTCACTTCGAGAGATTGCTCAACAACCATGCTGGCATCTGACCC	2245
	DB	1741	GTTGGGGAAGCTCTTTCACTTCGAGAGATTGCTCAACAACCATGCTGGCATCTGACCC	1800
	OY	2246	TCTTACTCTGATGATCCCTCTCTCTCCCGGTGGAGAGCTTACCCTTAAGTCCCGG	2305
	DB	1801	TCTTACTCTGATGATCCCTCTCTCTCCCGGTGGAGAGCTTACCCTTAAGTCCCGG	1860
	OY	2306	GATAGGCAAAAGTGGGCGGCGCGGAGACGCTGGGAGATTCCGGCGGACGCGACGC	2365
	DB	1861	GATAGGCAAAAGTGGGCGGCGCGGAGACGCTGGGAGATTCCGGCGGACGCGACGC	1920
	OY	2366	GCGCACCTGGAGACGGCGGCGCTGCTCCGAGAGCGCTTCCGGGCGCT	2425
	DB	1921	GCGCACCTGGAGACGGCGGCGCTGCTCCGAGAGCGCTTCCGGGCGCT	1980
	OY	2426	TAGGACCTTCCCTCAGGCGCGGCTCTCAAGCATGAGAGACTTCTTTCTGGGCT	2485
	DB	1981	TAGGACCTTCCCTCAGGCGCGGCTCTCAAGCATGAGAGACTTCTTTCTGGGCT	2040
	OY	2486	TGATCGGAAGTCTGATCCGCTGGGAGCGGCGGCGAGCTCCGGCGCTCTCGGAGCA	2545
	DB	2041	TGATCGGAAGTCTGATCCGCTGGGAGCGGCGGCGAGCTCCGGCGCTCTCGGAGCA	2600
	OY	2546	CTGCGCTCCAGCTTGAAGTGGCGCTGGGAGCGGAGCAGAAATTGAACGGAACTCTGGGA	2605
	DB	2101	CTGCGCTCCAGCTTGAAGTGGCGCTGGGAGCGGAGCAGAAATTGAACGGAACTCTGGGA	2160
	OY	2606	AGCTT 2610	
	DB	2161	AGCTT 2165	
	RESULT 7			
	LOCUS	AX026091	2380 bp	DNA linear PAT 16-SEP-2000
	DEFINITION	Sequence 3 from Patent DE19847779.		
	ACCESSION	AX026091		
	VERSION	AX026091.1 GI:10187522		
	KEYWORDS			
	SOURCE	Homo sapiens (human)		
	ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
	REFERENCE	1. Mueller-Schilling M., Kramerer, P. and Oren, M. Novel receptor data useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy Patent: DE 19847779-C 3 03-FEB-2000;		
	JOURNAL	DEUTSCHES KREBSFORSCH (DE) Location/Qualifiers		
	FEATURES	1. .2380 /organism="Homo sapiens" /mol type="genomic DNA" /db xref="taxon:9606"		
	BASE COUNT	579 a 568 g 638 t		
	ORIGIN			
	Query Match	59.1%; Score 1899.6; DB 6; Length 2380;		
	Best Local Similarity	84.9%; Pred. No. 0;		
	Matches 2348; Conservative	0; Mismatches 24; Indels 394; Gaps		
	448 AGCTTTTGGCTACATTTTATTGTAAGTAGTATAATCACTCATCTGACTG 507			
	1 ACCTTTTGGCTACATTTTATTGTAAGTAGTATAATCACTCATCTGACTG 60			

508 GGCATTAATGATAGTATTTAAGTAAAGAGATCCATATGAGTGTGCTTATPAT 567  
Db 61 GGCATTAATGATAGTATTTAAGTAAAGAGATCCATATGAGTGTGCTTATPAT 567  
Qy 568 TCACATCTAAGAGATGATGATTTGTCAATGTCCTTCCCTTTTCTCTTCCCT 120  
Db 121 TCACATCTAAGAGATGATGATTTGTCAATGTCCTTCCCTTTTCTCTTCCCT 627  
Qy 628 CTTTCATCTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 180  
Db 181 CTTTCATCTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 687  
Qy 688 CTTTCATCTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 240  
Db 241 CTTTCATCTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 747  
Qy 748 AAATTTGAGAGATATACAGAGATGCGCATATCCATCTCTTATCCCATCTTTT 300  
Db 301 AAATTTGAGAGATATACAGAGATGCGCATATCCATCTCTTATCCCATCTTTT 807  
Qy 808 TGTGCTATTTGATGCTCAGAGTGTGTGACAGAGGCTGACGCGCGAGGCTTCTCAT 360  
Db 361 TGTGCTATTTGATGCTCAGAGTGTGTGACAGAGGCTGACGCGCGAGGCTTCTCAT 867  
Qy 868 GGCATCTAAGCTCTACTGAAAGGTGAGACAGAGACAGCTTATCAACCTTACAGATG 420  
Db 421 GGCATCTAAGCTCTACTGAAAGGTGAGACAGAGACAGCTTATCAACCTTACAGATG 927  
Qy 928 GTGTAAAGTCAAGTACAGATGCAAAACAGAGGTATGAAAGGCTTACAGAGATG 480  
Db 481 GTGTAAAGTCAAGTACAGATGCAAAACAGAGGTATGAAAGGCTTACAGAGATG 987  
Qy 988 CTAACTTAATTTGAGGCTCC - AACAGGCTCCAGAGAAATGCTCACTGAGAGAGACC 540  
Db 541 CTAACTTAATTTGAGGCTCCAGAGAAATGCTCACTGAGAGAGACC 1046  
Qy 1047 TGAAGATGAACAGTGGCTAAGCAAGGCTTATTAATGTGTATTAATGAGTGAATCT 600  
Db 601 TGAAGATGAACAGTGGCTAAGCAAGGCTTATTAATGTGTATTAATGAGTGAATCT 1106  
Qy 1107 AATTGGAGAGAGAGAGTGTGACAGAGTGAAGGCTTGTGAGACATGCGAAAGG 660  
Db 661 AATTGGAGAGAGAGTGTGACAGAGTGAAGGCTTGTGAGACATGCGAAAGG 1166  
Qy 1167 AATPCTGAAACCTTTAGTGTGCTCAGTCTGAACTGATCAAAATTCAGGTCAGTAAATG 720  
Db 721 AATPCTGAAACCTTTAGTGTGCTCAGTCTGAACTGATCAAAATTCAGGTCAGTAAATG 1226  
Qy 1227 AATPCTGAAACCTTTAGTGTGCTCAGTCTGAACTGATCAAAATTCAGGTCAGTAAATG 780  
Db 781 AATPCTGAAACCTTTAGTGTGCTCAGTCTGAACTGATCAAAATTCAGGTCAGTAAATG 1286  
Qy 1287 CCGTCCAAAGCAATGATGCTTTGAAACAGTGTGACAGAGACAGAAAGATTAACAAGA 840  
Db 841 CCGTCCAAAGCAATGATGCTTTGAAACAGTGTGACAGAGACAGAAAGATTAACAAGA 1346  
Qy 1347 TTTTATTTTAAAGAAATTTGGCCAGAGAAATTAATGATTAACAGAGACAGAAAGATTAAGA 900  
Db 901 TTTTATTTTAAAGAAATTTGGCCAGAGAAATTAATGATTAACAGAGACAGAAAGATTAAGA 1406  
Qy 1407 GAATGTTAATATAGCTGGGCTATGCGATTTGCTTAAGTGTGATCTTTTCTTCTC 960  
Db 961 GAATGTTAATATAGCTGGGCTATGCGATTTGCTTAAGTGTGATCTTTTCTTCTC 1466  
Qy 1467 TTGAGAAATTAATACTGAGGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1020  
Db 1021 TTGAGAAATTAATACTGAGGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1526  
Qy 1527 TTTTCAATATGTTAATGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1080  
Db 1081 TTTTCAATATGTTAATGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1586  
Qy 1587 CAAACATGACAGCCAGTCAATATGCGCCAGAGTCTTTCTGAGTGAATCTCAAGCAATTA 1646

Db 1141 CAAACATGACAGCCAGTCAATATGCGCCAGAGTCTTTCTGAGTGAATCTCAAGCAATTA 1200  
Qy 1647 GCCAAGGCTCCCTGATACCCAGAGAGACCTTGCCTGCTGAGCTCCTTCTTCAAGAC 1706  
Db 1201 GCCAAGGCTCCCTGATACCCAGAGAGACCTTGCCTGCTGAGCTCCTTCTTCAAGAC 1260  
Qy 1707 CTCCCAACCTTCCAGAGTGAATCTACAGAGAGACCTTTTGAAGAGGACAGAGAGGCGGCT 1320  
Db 1261 CTCCCAACCTTCCAGAGTGAATCTACAGAGAGACCTTTTGAAGAGGACAGAGAGGCGGCT 1320  
Qy 1767 CTGAGGCTCCCTGATACCCAGAGAGACCTTGCCTGCTGAGCTCCTTCTTCAAGAC 1826  
Db 1321 CTGAGGCTCCCTGATACCCAGAGAGACCTTGCCTGCTGAGCTCCTTCTTCAAGAC 1380  
Qy 1827 TGCCTATTTGTCAGAGAACCTTGAATCTCTTCTTCAAGCTTCTTCCCTTCCCTTCCCT 1886  
Db 1381 TGCCTATTTGTCAGAGAACCTTGAATCTCTTCTTCAAGCTTCTTCCCTTCCCTTCCCT 1440  
Qy 1887 CCGCGCAGAGCCAGTGTGTAATCAATGAGACCTTCCCAACCCGAGGCTTCCCAAGC 1946  
Db 1441 CCGCGCAGAGCCAGTGTGTAATCAATGAGACCTTCCCAACCCGAGGCTTCCCAAGC 1500  
Qy 1947 AGGCTTCCCTTCCATCTCTCTGACACCGAGGCTTTTGTGAGCTGCTCTGATCTGCG 2006  
Db 1501 AGGCTTCCCTTCCATCTCTCTGACACCGAGGCTTTTGTGAGCTGCTCTGATCTGCG 1560  
Qy 2007 GCAAGATGACACAGAGTGTCAAAAGACCTTGTGAGAGTGAAGGAGCGGTTTACAA 2066  
Db 1561 GCAAGATGACACAGAGTGTCAAAAGACCTTGTGAGAGTGAAGGAGCGGTTTACAA 1620  
Qy 2067 GTGACTTGGCTGAGGCTTCAAGGCGGAGCACTGACAGAGACCAACCTTGAAGGCTTCAAG 2126  
Db 1621 GTGACTTGGCTGAGGCTTCAAGGCGGAGCACTGACAGAGACCAACCTTGAAGGCTTCAAG 1680  
Qy 2127 TGGCTGCGCAGGCGAGCTGCTTCTTCTGCGGAGGCTTGTGAGACCTTCAAGGAG 2186  
Db 1681 TGGCTGCGCAGGCGAGCTGCTTCTTCTGCGGAGGCTTGTGAGACCTTCAAGGAG 1727  
Qy 2187 TGGGAGAGGCTTCTTCACTTGGAGAGATTTGCTCAACCAATGCTGGGCTTGAACCTT 2246  
Db 1728 TGGGAGAGGCTTCTTCACTTGGAGAGATTTGCTCAACCAATGCTGGGCTTGAACCTT 1740  
Qy 2247 CTTACTTGTGATTCCTTCTTCTGCGGAGTGAAGGCTTACCCCTTATGTCGCGGG 2306  
Db 1741 CTTACTTGTGATTCCTTCTTCTGCGGAGTGAAGGCTTACCCCTTATGTCGCGGG 1740  
Qy 2307 ATGAGCAAGTGGGCGGAGCGGAGACGCTGCGGATTTGCGGCGAGCGGCGACGCG 2366  
Db 1741 ATGAGCAAGTGGGCGGAGCGGAGACGCTGCGGATTTGCGGCGAGCGGCGACGCG 1765  
Qy 2367 GGCACCTGGAGGCGGAGCTGCTGCGGAGGCTTGAAGCTGCTCCCGGAGGCTGTT 2426  
Db 1766 GGCACCTGGAGGCGGAGCTGCTGCGGAGGCTTGAAGCTGCTCCCGGAGGCTGTT 1765  
Qy 2427 AGGACTTCTCAGGCGCGGCTGCTCAAGAACATGAGAGACTTCTTCTTGGGCTT 2486  
Db 1766 AGGACTTCTCAGGCGCGGCTGCTCAAGAACATGAGAGACTTCTTCTTGGGCTT 1765  
Qy 2487 GATGCAAGTGTGATTCGCGTGGAGAGCGGAGAGCTCCGCGCTCTCTGAGAGACAC 2546  
Db 1766 GATGCAAGTGTGATTCGCGTGGAGAGCGGAGAGCTCCGCGCTCTCTGAGAGACAC 1765  
Qy 2547 TGCCTCAGCTTATGATGCGCTGCGGAGCGGAGACAGAAATTAAGCGGAAATCTGGGAA 2606  
Db 1766 TGCCTCAGCTTATGATGCGCTGCGGAGCGGAGACAGAAATTAAGCGGAAATCTGGGAA 1767  
Qy 2607 GCTTTAGGCTGCTGAGAGGAGAGCCCGGCTTTGAGAGAGAGAGAGAGCTTGAAGCAAG 2666  
Db 1768 GCTTTAGGCTGCTGAGAGGAGAGCCCGGCTTTGAGAGAGAGAGAGAGCTTGAAGCAAG 1827  
Qy 2667 CTTGACAGAGCAAGCCAAAGTCTCGGCGCGGAGTGAATGCGGCGCGCGCGCG 2726

```

Db      1828 CCTGACGAGCCAAAGGCTCCGCGCGGGTGGTGAAGTCCGCCGCCGCCGCC 1887
Oy      2727 GGGGGGGGGAGAGAGCCCTACAGCCTTACAGAACATATTTCTCATTTTCTGGAGTTCT 2786
Oy      1888 GGGGGGGGGAGAGAGCCCTGAGCCTTACAGAACATATTTCTCATTTTCTGGAGTTCT 1947
Oy      2787 CAGAGTGAAGAAATTAAGTACAGACCGAGACGATGTTAAGCCGGAGGCTCGGAGAAACG 2846
Db      1948 CAGAGCTGAGAAATTAAGTACAGACCGAGACGATGTTAAGCCGGAGGCTCGGAGAAACG 2007
Oy      2847 GCACTTTTCTTTCTGCAAAAAGTTATATGAGGAGCTGAATGAGCTTCTGAGAGCTTGT 2906
Oy      2008 GCACCTTTTCTTTCTGCAAAAAGTTATATGAGGAGCTGAATGAGCTTCTGAGAGCTTGT 2067
Oy      2907 ACCGTTTTTATTTGTCACACAGAAAAGAACTGCTTCTCCCTCCGGGAATTTCTCT 2966
Db      2068 ACCGTTTTTATTTGTCACACAGAAAAGAACTGCTTCTCCCTCCGGGAATTTCTCT 2127
Oy      2967 CTTTAAAGCTGTAAGTGGCTGCTGAGTGGTTCAATTTGTTTGTCTTCTGCTCTCT 3026
Db      2128 CTTTAAAGCTGTAAGTGGCTGCTGAGTGGTTCAATTTGTTTGTCTTCTGCTCTCT 2187
Oy      3027 CTTTCTTTTGGCTTTCTTTCTTTGCTTCACTCCCATGCTGATTTCTGCTTGTCTCT 3086
Db      2188 CTTTCTTTTGGCTTTCTTTCTTTGCTTCACTCCCATGCTGATTTCTGCTTGTCTCT 2247
Oy      3087 CTGGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3146
Db      2248 CTGGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2307
Oy      3147 ACTGAGAGAGCTGTTTGAAGAGTCCCTGCTCGTCAAGATGCTGAGTTGAGATGCT 3206
Db      2308 ACTGAGAGAGCTGTTTGAAGAGTCCCTGCTCGTCAAGATGCTGAGTTGAGATGCT 2367
Oy      3207 TCAAG 3212
Oy      2368 TCAAG 2373
Db      3207 TCAAG 3212
Db      2368 TCAAG 2373

RESULT 8
LOCUS   HSAPT1 1877 bp DNA linear PRI 08-JUN-1995
DEFINITION H.sapiens APT gene, exon 1.
ACCESSION X81335
VERSION X81335.1 GI:537410
KEYWORDS APO-1 gene; apt gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Chordata; Craniata; Vertebrata; Euteleostomi;
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
  1 Behrmann, I., Walczak, H. and Kramer, P. H.
    TITLE Structure of the human APO-1 gene
    JOURNAL Eur. J. Immunol. 24 (12), 3057-3062 (1994)
MEDLINE 95104292
PUBMED 7528667
REFERENCE
  2 (bases 1 to 1877)
    Krammer, P. H.
    TITLE Direct Submission
    JOURNAL Submitted (03-SEP-1994) P. H. Krammer, German Cancer Research
    Center, Tumorimmunology Programm, Im Neuenheimer Feld 280, 69120
    Heidelberg, FRG

FEATURES
  source
    1..1877
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="10"
    /map="10q23"
    /clone="CAPO-1"
    /clone_1ib="cosmid, pcos2EMBL"
    151..361
    /note="pyrimidine rich region"

misc_feature
  /note="pyrimidine rich region"

```

*Missing Exon 1 and intron*

```

gene
  join(1653..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  join(1653..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA1"
  1653..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA1
  /label=ex1a
  join(1670..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA2"
  1670..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA2
  /label=ex1b
  join(1689..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA3"
  1689..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA3
  /label=ex1c
  join(1704..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA4"
  1704..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA4
  /label=ex1d
  join(1707..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA5"
  1707..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA5
  /label=ex1e
  join(1708..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA6"
  1708..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA6
  /label=ex1f
  join(1709..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA7"
  1709..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA7
  /label=ex1g
  join(1710..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA8"
  1710..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA8
  /label=ex1h
  join(1711..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA9"
  1711..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA9
  /label=ex1i
  join(1712..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA10"
  1712..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA10
  /label=ex1j
  join(1713..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA11"
  1713..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA11
  /label=ex1k
  join(1714..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA12"
  1714..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA12
  /label=ex1l
  join(1715..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA13"
  1715..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA13
  /label=ex1m
  join(1716..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA14"
  1716..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA14
  /label=ex1n
  join(1717..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA15"
  1717..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA15
  /label=ex1o
  join(1718..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA16"
  1718..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA16
  /label=ex1p
  join(1719..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA17"
  1719..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA17
  /label=ex1q
  join(1720..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA18"
  1720..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA18
  /label=ex1r
  join(1721..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA19"
  1721..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA19
  /label=ex1s
  join(1722..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA20"
  1722..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA20
  /label=ex1t
  join(1723..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA21"
  1723..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA21
  /label=ex1u
  join(1724..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA22"
  1724..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA22
  /label=ex1v
  join(1725..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA23"
  1725..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA23
  /label=ex1w
  join(1726..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA24"
  1726..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA24
  /label=ex1x
  join(1727..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA25"
  1727..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA25
  /label=ex1y
  join(1728..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA26"
  1728..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA26
  /label=ex1z
  join(1729..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA27"
  1729..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA27
  /label=ex1aa
  join(1730..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA28"
  1730..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA28
  /label=ex1ab
  join(1731..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA29"
  1731..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA29
  /label=ex1ac
  join(1732..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA30"
  1732..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA30
  /label=ex1ad
  join(1733..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA31"
  1733..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA31
  /label=ex1ae
  join(1734..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA32"
  1734..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA32
  /label=ex1af
  join(1735..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA33"
  1735..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA33
  /label=ex1ag
  join(1736..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA34"
  1736..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA34
  /label=ex1ah
  join(1737..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA35"
  1737..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA35
  /label=ex1ai
  join(1738..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA36"
  1738..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA36
  /label=ex1aj
  join(1739..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA37"
  1739..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA37
  /label=ex1ak
  join(1740..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA38"
  1740..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA38
  /label=ex1al
  join(1741..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA39"
  1741..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA39
  /label=ex1am
  join(1742..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA40"
  1742..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA40
  /label=ex1an
  join(1743..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA41"
  1743..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA41
  /label=ex1ao
  join(1744..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA42"
  1744..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA42
  /label=ex1ap
  join(1745..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA43"
  1745..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA43
  /label=ex1aq
  join(1746..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA44"
  1746..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA44
  /label=ex1ar
  join(1747..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA45"
  1747..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA45
  /label=ex1as
  join(1748..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA46"
  1748..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA46
  /label=ex1at
  join(1749..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA47"
  1749..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA47
  /label=ex1au
  join(1750..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA48"
  1750..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA48
  /label=ex1av
  join(1751..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA49"
  1751..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA49
  /label=ex1aw
  join(1752..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA50"
  1752..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA50
  /label=ex1ax
  join(1753..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA51"
  1753..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HemRNA51
  /label=ex1ay
  join(1754..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..828,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HemRNA52"
  1754..1810
  /gene="APT"
```

[illegible]

Missing Exon 1 and intron

DEFINITION	H. sapiens Fas, Apo-1 gene (promoter and exon I).	1608 bp	DNA	linear	PRI 17-FEB-1995
ACCESSION	X82279				
VERSION	X82279.1	GI:673405			
KEYWORDS	Apo-1 gene; Fas gene.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1				
AUTHORS	Cheng, J., Liu, C., Koopman, W. J. and Mountz, J. D.				

TITLE Characterization of human Fas gene. Exon/Intron organization and  
 JOURNAL Promoter region  
 MEDLINE J. Immunol. 154 (3), 1239-1245 (1995)  
 PUBMED 95123075  
 REFERENCE 7529798  
 AUTHORS 2 (bases 1 to 1608)  
 TITLE Cheng J.  
 JOURNAL Direct Submission  
 COMMENT Submitted (20-OCT-1994) J. Cheng, Univ. of Alabama at Birmingham,  
 Birmingham, AL 35294-0007, USA  
 Related sequences: M67454 and X63717.  
 FEATURES  
 source Location/Qualifiers  
 1.1608  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /map="10q23 or 10q24.1"  
 /issue\_type="placenta"  
 /clone\_id="Placental genomic DNA"  
 /gene="Fas, Apo-1"  
 /gene="Fas, Apo-1"  
 /gene="Fas, Apo-1"  
 BASE COUNT 398 a 421 c 423 g 366 t  
 ORIGIN  
 Query Match 47.8%; Score 1536.8; DB 9; Length 1608;  
 Best Local Similarity 98.6%; Pred. No. 0;  
 Matches 1582; Conservative 0; Mismatches 17; Indels 5; Gaps 3;  
 gene  
 exon  
 767 AGAGATGCCATATACCATCTCTCTATCCCATCTCTTTTGTGTCTATTAAGATCTCA 826  
 7 AAGAGATGCCATATACCATCTCTCTATCCCATCTCTTTTGTGTCTATTAAGATCTCA 66  
 827 GAGTGTGCACAGAGCTGCGACGCCAGGCTCTTCCATGACGACTAACTTACTGA 886  
 67 GAGTGTGCACAGAGCTGCGACGCCAGGCTCTTCCATGACGACTAACTTACTGA 126  
 887 AAGTGTGCACAGAGCTGCGACGCCAGGCTCTTCCATGACGACTAACTTACTGA 946  
 127 AAGTGTGCACAGAGCTGCGACGCCAGGCTCTTCCATGACGACTAACTTACTGA 186  
 947 ATGCAAAACACAGAGCTGCGACGCCAGGCTCTTCCATGACGACTAACTTACTGA 1006  
 187 ATGCAAAACACAGAGCTGCGACGCCAGGCTCTTCCATGACGACTAACTTACTGA 246  
 1007 CC-AACAGCTCCAGAGAAATGTCACTGAGAGAGAGCTGAGATGAACAGTGGC 1065  
 247 CCAACAGCTCCAGAGAAATGTCACTGAGAGAGAGCTGAGATGAACAGTGGC 306  
 1066 TAAGCAAGAGCTTATTAATGTGTTTATTAATGCTTGAATCTTAATGGGAAGGAGAGG 1125  
 307 TAAGCAAGAGCTTATTAATGTGTTTATTAATGCTTGAATCTTAATGGGAAGGAGAGG 366  
 1126 TTGCAAGTGTGAGGTCAGAGCTTGTGTCAGATGCCAAGAGATCTGAACCTTTAGT 1185  
 367 TTGCAAGTGTGAGGTCAGAGCTTGTGTCAGATGCCAAGAGATCTGAACCTTTAGT 426  
 1186 TGTCCAGTGTGAGGTCAGAGCTTGTGTCAGATGCCAAGAGATCTGAACCTTTAGT 1245  
 427 TGTCCAGTGTGAGGTCAGAGCTTGTGTCAGATGCCAAGAGATCTGAACCTTTAGT 486  
 1246 CTTTCTGTAAATTCATGCTTAACCTTAAGAGCTTACCGTCCAAAGCAATAGTG 1305  
 487 CTTTCTGTAAATTCATGCTTAACCTTAAGAGCTTACCGTCCAAAGCAATAGTG 546  
 1306 ACTTTGAACAGTGTCCACAGAGCAAGAAATTAAGATTTTAAAGAAAT 1365  
 547 ACTTTGAACAGTGTCCACAGAGCAAGAAATTAAGATTTTAAAGAAAT 606  
 1366 GCGCAGAAATTAAGATTAAGAGCAAGAAATTAAGATTTTAAAGAAAT 1425

607 GCGCAGAAATTAAGATTAAGAGCAAGAAATTAAGATTTTAAAGAAAT 666  
 1426 GCGTATGCAATTTGGCTTAAGTGTGTTTCTCTTGAAGATTAAGAAATTAAG 1485  
 667 GCGTATGCAATTTGGCTTAAGTGTGTTTCTCTTGAAGATTAAGAAATTAAG 726  
 1486 GGGCCCTCCCTTTTCAAGAGCTTATGAGGCAATCTGTAATTTTCAATAGTTAAGT 1545  
 727 GGGCCCTCCCTTTTCAAGAGCTTATGAGGCAATCTGTAATTTTCAATAGTTAAGT 786  
 1546 TCCATTCAGAAACGCTGTGAGCTTCAATGTGAGGCAACAAATGAGCAAGTCCAGT 846  
 787 TCCATTCAGAAACGCTGTGAGCTTCAATGTGAGGCAACAAATGAGCAAGTCCAGT 846  
 1606 AAATGCCCGCAAGCTTCTGAGTGTGAGTCCAGCAATTAAGCAAGTCCCTGTAACCA 1665  
 847 AAATGCCCGCAAGCTTCTGAGTGTGAGTCCAGCAATTAAGCAAGTCCCTGTAACCA 906  
 1666 GCGCAGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAG 1725  
 907 GCGCAGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAG 966  
 1726 GAACTACAGAGAGCTTGTGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 1785  
 967 GAACTACAGAGAGCTTGTGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 1845  
 1786 GTGAGCAGTCCAGCACTGAGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 1845  
 1027 GTGAG-ATGCCAGCCTGAGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 1885  
 1846 CCTGAGCTCCTGCTGAGCCTGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAG 1905  
 1086 CCTGAGCTCCTGCTGAGCCTGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAG 1145  
 1906 TGAATCAATGAGAGCTTGTGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 1965  
 1146 TGAATCAATGAGAGCTTGTGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 1205  
 1966 CTGACCAAGAGGCTTGTGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 2025  
 1206 CTGACCAAGAGGCTTGTGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 1265  
 2026 GTTCAAGAGCTTGTGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 2085  
 1266 GTTCAAGAGCTTGTGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 1325  
 2086 AAGGCGGCGCACTGAGCAAGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 2145  
 1326 AAGGCGGCGCACTGAGCAAGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 1385  
 2146 GCTCTTCTCCGCGGCTTGTGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 2205  
 1386 GCTCTTCTCCGCGGCTTGTGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 1445  
 2206 TCGAGAGATTTGCAACACATGCTGAGGCAAGAGGCGGCTTGTGAGCTTGTGAG 2265  
 1446 TCGAGAGATTTGCAACACATGCTGAGGCAAGAGGCGGCTTGTGAGCTTGTGAG 1505  
 2266 CTCCTGCGGCGGCTTGTGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 2325  
 1506 CTCCTGCGGCGGCTTGTGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 1565  
 2326 GCGGCGGCGGCTTGTGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 2365  
 1566 GCGGCGGCGGCTTGTGAGGCAAGAGGCGGCTTGTGAGCTTGTGAGCTTGTGAG 1606

RESULT 10  
 AX347321/C 1608 bp DNA linear PAT 01-FEB-2002  
 LOCUS AX347321  
 DEFINITION Sequence 2392 from Patent WO0200928.  
 ACCESSION AX347321  
 VERSION AX347321.1 GI:18495209

KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
SOURCE

synthetic construct  
synthetic construct  
artificial sequences.

Olek, A., Piepenbrock, C. and Berlin, K.  
Diagnosis of diseases associated with the immune system  
Patent: WO 0200928-A 2392 03-JAN-2002;  
Epigenomics AG (DE)

Location/Qualifiers  
1..1608  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 366 a 51 c 421 g 770 t

ORIGIN

Query Match 29.9%; Score 961; DB 6; Length 1608;  
Best Local Similarity 76.6%; Pred. No. 1.2e-217;  
Matches 1202; Conservative 0; Mismatches 365; Indels 2; Gaps 2;

767 AAGAGATGCCATATACATCCCTTATCCCACTTTCTTTTGTGTATAGATGCTCA 826  
Db 1602 AAAAAAATACCTATATACATCTCTTATCCCACTTTCTTTTGTATATTAATCTCA 1543

QY 827 GAGTGTGACAAAGCTGTGCGCGCCAGGGCTTCTCATAGGCACTAAGCTACTGA 886  
Db 1542 AATATATACAAAGCTGTGCGCGCCAGGGCTTCTCATAGGCACTAAGCTACTGA 1483

QY 887 AAGGTGAACAGACAGCTTACACCTTACAGCTGAGTGTGAGTGAAGTGAAG 946  
Db 1482 AAAAAAATACCTATATACATCCCTTATCCCACTTTCTTTTGTATATTAATCTCA 1423

QY 947 ATGCAAAACAGGGTGTATGAAAGCCCTCAAGAGGGTAACTTAAGTTGGAGGC 1006  
Db 1422 ATCAAAACAGGGTGTATGAAAGCCCTCAAGAGGGTAACTTAAGTTGGAGGC 1363

QY 1007 CC-AACAGGCTCAGAGAAATGTCACTGAGAGAGAGCTTAAGATGAAGTGAAG 1065  
Db 1362 CCAACAGGCTCAGAGAAATGTCACTGAGAGAGAGCTTAAGATGAAGTGAAG 1303

QY 1066 TAAGCAAGGGTATATATGTGTATTAATGGGTGAATCTAATTTGGAAGGAGAGAG 1125  
Db 1302 TAACCAAGGGTATATATGTGTATTAATGGGTGAATCTAATTTGGAAGGAGAGAG 1243

QY 1126 TTGCAGAGTGTGAGAGCTGTGGAGAGTGCAGAAAGAACTGAAACCTTTAGTG 1185  
Db 1242 TTACCAAGTGTGAGAGCTGTGGAGAGTGCAGAAAGAACTGAAACCTTTAGTG 1183

QY 1186 TGTCAAGTGTGAGAGCTGTGGAGAGTGCAGAAAGAACTGAAACCTTTAGTG 1245  
Db 1182 TATCAAGTGTGAGAGCTGTGGAGAGTGCAGAAAGAACTGAAACCTTTAGTG 1123

QY 1246 CTTTGTGAAATTCATGCTAACTAAGAGTATCCGTTCCAAAGCAATAGTG 1305  
Db 1122 CTTTGTGAAATTCATGCTAACTAAGAGTATCCGTTCCAAAGCAATAGTG 1063

QY 1306 ACTTGAACAGTGTGAGAGCTGTGGAGAGTGCAGAAAGAACTTATTAAGAAAT 1365  
Db 1062 ACTTGAACAGTGTGAGAGCTGTGGAGAGTGCAGAAAGAACTTATTAAGAAAT 1003

QY 1366 GGGCAGAAATATAGTAAAGAGAGAGTATTTGTGAATGTTTAAATAGCTGG 1425  
Db 1002 AACCAAAATATATTAATTAAGAGAGAGTATTTGTGAATGTTTAAATAGCTGA 943

QY 1426 GGGTATGAGATTTGGCTTATGCTTTTCTCTTGAAGAAATTAAGTAAAG 1485  
Db 942 AACTATGAGATTTGGCTTATGCTTTTCTCTTGAAGAAATTAAGTAAAG 883

QY 1486 GGGCTCTCTCTTTCAGAGCTTATGCGCAACATGCACTTTTTCATAGGTTAGTG 1545  
Db 882 AAGCTCTCTCTTTCAGAGCTTATGCGCAACATGCACTTTTTCATAGGTTAGTG 823

QY 1546 TCATTCAGAAAGCTGTGTAGAGCTTCTATGTCAGACCAACATGAGAGCCCATC 1605  
Db 822 TCATTCAGAAAGCTGTGTATTAAGCTTCTATGTCAGACCAACATGAGAGCCCATC 763

QY 1606 AATGCCCCGCAAGCTTCTGTGAGTACCCAGATTTAGCAAGGCTCTGTACCA 1665  
Db 762 AATGCCCCGCAAGCTTCTGTATATATGTCAGACCAACATGAGAGCCCATC 703

QY 1666 GGCAGAGCTCTGGGCTGTAGCTTCATTTCTCTTCAAGAGCTCCCAACTGCCAGTT 1725  
Db 702 AAGCAAACTCTAGCTCTTAACTTCATCTCTTCAAAAGCTCCCAACTGCCAGTT 643

QY 1726 GAATCAAGCAGAAAGCTTTAGAAAGGAGAGGCGGCTCTGAGAGTCTCACTGA 1785  
Db 642 AACTCAACACAAAGCTTTTAAAGAAACAAAGAGGAGTCTCACTCACTGA 583

QY 1786 GTGAGCATGCGCAGCATGTGAGAGAGCCCGGAGAGAGATGCCATTTGAGCAAG 1845  
Db 582 AT-AAGATCAACACCTTAAAGAGAGCCCGGAGAGAGATGCCATTTGAGCAAG 524

QY 1846 CCTGACTCTTCTGACCTGACTTCTGCTCCCTCCCTTACCGGCGCAGGCGCAAGTTGC 1905  
Db 523 CCTTAACTCTTCTGACCTTAACTTCTCCCTCCCTTACCGGCGCAGGCGCAAGTTGC 464

QY 1906 TGAATCAATGAGAGCTCTCCCAAGCCGAGGCTTCCCAAGAGAGGCTTCCCATCTC 1965  
Db 463 TGAATCAATGAGAGCTCTCCCAAGCCGAGGCTTCCCAAGAGAGGCTTCCCATCTC 404

QY 1966 CTGACCAAGGAGCTTCTGAGAGCTGCTCTGATCTCTGAGAGAGTGAACAGAGT 2025  
Db 403 CTGACCAAGGAGCTTCTGAGAGCTGCTCTGATCTCTGAGAGAGTGAACAGAGT 344

QY 2026 GTTCAAGAGCTTCTGAGAGAGTGAAGAGAGGCTTTAGAGTACTTGTGAGAGCTC 2085  
Db 343 GTTCAAGAGCTTCTGAGAGAGTGAAGAGAGGCTTTAGAGTACTTGTGAGAGCTC 284

QY 2086 AGGGCGGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2145  
Db 283 AGGGCGGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 224

QY 2146 GCTCTCTCTCCGCGGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2205  
Db 223 GCTCTCTCTCCGCGGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 164

QY 2206 TCGAGAGATTTCTCAACACATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 2265  
Db 163 TCGAGAGATTTCTCAACACATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 104

QY 2266 CTCTGCGGAG 2325  
Db 103 CTCTGCGGAG 44

QY 2326 CCGGAGAG 2334  
Db 43 CCGGAGAG 35

RESULT 11  
AX347320 1608 bp DNA PAT 01-FEB-2002  
LOCUS AX347320  
DEFINITION Sequence 2391 from Patent WO0200928.  
ACCESSION AX347320  
VERSION AX347320.1 GI:18495208  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

synthetic construct  
synthetic construct  
artificial sequences.

Olek, A., Piepenbrock, C. and Berlin, K.  
Diagnosis of diseases associated with the immune system  
Patent: WO 0200928-A 2391 03-JAN-2002;  
Epigenomics AG (DE)





Query Match 22.2% Score 713; DB 6; Length 720;  
 Best Local Similarity 100.0%; Pred. No. 1,3e-158;  
 Matches 713; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2500 GATCCCGCTGGGAGGCGGCGGAGCTCCGCGCTCTCTGAGAGACACACTGCGGCTCCACGTT 2559  
 DB 1 GATCCCGCTGGGAGGCGGCGGAGCTCCGCGCTCTCTGAGAGACACACTGCGGCTCCACGTT 2559  
 QY 2560 GAGGTGGCGCTGGGCGGCGGAGCAAGATTGAACCGGAAGTCTGGGAAGCTTTAGGTCGC 2619  
 DB 61 GAGGTGGCGCTGGGCGGCGGAGCAAGATTGAACCGGAAGTCTGGGAAGCTTTAGGTCGC 2619  
 QY 2620 TGGAGGGGAGACCCCGGCTTGGAGAGAGAGCGGAATCTCTGACAAGCCCTTACAGGCA 2679  
 DB 121 TGGAGGGGAGACCCCGGCTTGGAGAGAGAGCGGAATCTCTGACAAGCCCTTACAGGCA 2679  
 QY 2680 GCCAAGGTCCGCTCCGCGGCGGCGGAGTGTGAGTCCGCGCGCCCGCGGCGGCGGAGGA 2739  
 DB 181 GCCAAGGTCCGCTCCGCGGCGGCGGAGTGTGAGTCCGCGCGCCCGCGGCGGCGGAGGA 2739  
 QY 2740 GAGCCTACAGCCTCAGAACACATATGCTCATTTCTGCAATTTCTCAGAGTAGA 2799  
 DB 241 GAGCCTACAGCCTCAGAACACATATGCTCATTTCTGCAATTTCTCAGAGTAGA 2799  
 QY 2800 TAAGTCAGCACCGAAGAGTGTGTAAGCCGAGAGGCTTCGAGAGACGCGACCTTTCTT 360  
 DB 301 TAAGTCAGCACCGAAGAGTGTGTAAGCCGAGAGGCTTCGAGAGACGCGACCTTTCTT 360  
 QY 2860 CTCGAAAAAGTTATATGAGGCGCTGAATGAGCTTCTGAGAGCTTTGTTTACCTTTTATT 2919  
 DB 361 CTCGAAAAAGTTATATGAGGCGCTGAATGAGCTTCTGAGAGCTTTGTTTACCTTTTATT 2919  
 QY 2920 CTCGAAAAAGTTATATGAGGCGCTGAATGAGCTTCTGAGAGCTTTGTTTACCTTTTATT 420  
 DB 421 CTCGAAAAAGTTATATGAGGCGCTGAATGAGCTTCTGAGAGCTTTGTTTACCTTTTATT 2919  
 QY 2960 AGTGGCTGCTGAGTGTGTTTATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
 DB 481 AGTGGCTGCTGAGTGTGTTTATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3039  
 QY 3040 CCGCTTCTGAGCTGCACTCCATGCTGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 DB 541 CCGCTTCTGAGCTGCACTCCATGCTGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3039  
 QY 3100 TACTGCTTCCACCGACAGAACCCGCGGCTTATTTGTCGCAAGAACTTGAAGAGCTT 600  
 DB 601 TACTGCTTCCACCGACAGAACCCGCGGCTTATTTGTCGCAAGAACTTGAAGAGCTT 3159  
 QY 3160 GTTTTGAAGAGTCCCTGCTGAGAAATCCAGCTTGCAGATGCTTATCAAG 3212  
 DB 661 GTTTTGAAGAGTCCCTGCTGAGAAATCCAGCTTGCAGATGCTTATCAAG 713

RESULT 13  
 HSA279012 702 bp DNA linear PRI 02-APR-2001  
 LOCUS HSA279012  
 DEFINITION Homo sapiens partial CD95 gene for CD95 antigen (Apo-1 Fas), exon  
 VERSION A0279012.1 GI:13539240  
 KEYWORDS Apo-1 Fas; CD95 antigen; CD95 gene.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 Muschen, M., Re, D., Jungnickel, B., Diehl, V., Rajewsky, K. and  
 Kuppers, R. Somatic mutations of the CD95 gene in human B cells as a  
 TITLE 'Somatic mutations of the CD95 gene in human B cells as a  
 JOURNAL Unpublished  
 REFERENCE 2.

AUTHORS Muschen, M., Re, D., Brauminger, A., Wolf, J., Hanemann, M.L., Diehl, V.,  
 Kuppers, R. and Rajewsky, K.  
 TITLE Somatic mutations of the CD95 gene in Hodgkin- and Reed-Sternberg  
 cells  
 JOURNAL Unpublished  
 REFERENCE 3 (bases 1 to 702)  
 AUTHORS Muschen, M.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-SEP-2000) Muschen M., Department of Immunology,  
 Institute for Genetics, LFI B4 R705, Joseph-Stelzmann-Str. 9, 50931  
 Koeln, GERMANY  
 COMMENT Related sequences: D31968 X89101 A0279011 A0279013.  
 FEATURES  
 source  
 location/Qualifiers  
 1..702  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /cell\_type="Reed-Sternberg cells"  
 /note="Hodgkin's disease tissue-polymorphic allele (+275)"  
 /gene="CD95"  
 <1..313  
 /gene="CD95"  
 /standard\_name="Apo-1 Fas"  
 /number=1  
 <1..281  
 /gene="CD95"  
 /standard\_name="Apo-1 Fas"  
 /number=1  
 282..313  
 /standard\_name="Apo-1 Fas"  
 /number=1  
 /gene="CD95"  
 /codon\_start=1  
 /product="CD95 antigen"  
 /protein\_id="CAC35540.1"  
 /db\_xref="GI:13539241"  
 /translation="MLGIVTLPLV"  
 314..702  
 /gene="CD95"  
 /standard\_name="Apo-1 Fas"  
 /number=1

BASE COUNT 107 a 197 c 264 g 134 t  
 ORIGIN

Query Match 21.8% Score 698.8; DB 9; Length 702;  
 Best Local Similarity 99.7%; Pred. No. 3.1e-155;  
 Matches 700; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1946 GAGCTTCTCCATCCATCCTGACACACCGGCGCTTTGTGAGCTCGTCTGATCTCG 2005  
 DB 1 GAGCTTCTCCATCCATCCTGACACACCGGCGCTTTGTGAGCTCGTCTGATCTCG 2005  
 QY 2006 CGAAGAGTACACACAGAGTGTTCAAAGAGCTTCTGCGGAGTGTGAGGAGCGGTTTACG 2065  
 DB 61 CGAAGAGTACACACAGAGTGTTCAAAGAGCTTCTGCGGAGTGTGAGGAGCGGTTTACG 2065  
 QY 2066 AGTGAAGTGTGAGAGCTCAGAGGCGGCGCACTGCAAGAACACACCTGAGGCGGCGC 2125  
 DB 121 AGTGAAGTGTGAGAGCTCAGAGGCGGCGCACTGCAAGAACACACCTGAGGCGGCGC 2125  
 QY 2126 CTGAGCTGCGGCGGAGAGCTGCTTCTCCGCGGAGTGTGAGAGCGGCTCAGTACGGA 2185  
 DB 181 CTGAGCTGCGGCGGAGAGCTGCTTCTCCGCGGAGTGTGAGAGCGGCTCAGTACGGA 2185  
 QY 2186 GTTGGGGAAGCTTTTCACTTGGAGATGCTCAACACATGCTGGGATCTGGAGCC 2245  
 DB 241 GTTGGGGAAGCTTTTCACTTGGAGATGCTCAACACATGCTGGGATCTGGAGCC 2245  
 QY 2246 TCTTACCTGTGAGTCCCTCTCTCCGCGGAGTGTGAGAGCTTACCGGCTTATGTCGGG 2305  
 DB 301 TCTTACCTGTGAGTCCCTCTCTCCGCGGAGTGTGAGAGCTTACCGGCTTATGTCGGG 2305  
 QY 2306 GATAGGCAAGTGGGCGGCGGAGCGGCTGCGGAGTTTCCGCGGAGCGGCGGCGGCGGCGG 2365



Db 361 GATAGGCAAAAGTGGGGCGCGCGGAGACGCTGCGGATTGCGCGGACGCGCACCC 420

QY 2366 GGGGACCTTGGAGCGCGCGGCTGCTGGGGGAGGCGTTTGAAGACTGAGCTCCCGGGGCTGT 2425

Db 421 GGGGACCTTGGAGCGCGCGGCTGCTGGGGGAGGCGTTTGAAGACTGAGCTCCCGGGGCTGT 480

QY 2426 TAGGACCTTCCCTCAGCGCCGGGTGCTCAGAACGATGAGAGACTTTTCTTGGGCT 2485

Db 481 TAGGACCTTCCCTCAGCGCCGGGTGCTCAGAACGATGAGAGACTTTTCTTGGGCT 540

QY 2486 TGATGCGAAGTGTGATCCCGCTGGGCGAGCGGGGAGCTCCGGGCTCTCGAGAGCA 2545

Db 541 TGATGCGAAGTGTGATCCCGCTGGGCGAGCGGGGAGCTCCGGGCTCTCGAGAGCA 600

QY 2546 CTGCGCTCCACGTTGATGAGTGGCGGTGGGGGCGGACAGAAATTGAAGCGGAGCTCGGA 2605

Db 601 CTGCGCTCCACGTTGATGAGTGGCGGTGGGGGCGGACAGAAATTGAAGCGGAGCTCGGA 660

QY 2606 AGCTTTAGGCTGCTGAGAGGGGAGCCCGGTTGAGAGAGGA 2647

Db 661 AGCTTTAGGCTGCTGAGAGGGGAGCCCGGTTGAGAGAGGA 702

RESULT 14  
HSA279011 702 bp DNA linear PRI 02-APR-2001  
LOCUS Homo sapiens partial CD95 gene for CD95 antigen (Apo-1 Fas), exon  
DEFINITION 1.  
ACCESSION AJ279011 GI:13539238  
VERSION Apo-1 Fas; CD95 antigen; CD95 gene.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 Muschen, M., Re, D., Jungnickel, B., Diehl, V., Rajewsky, K. and Kuper, R. Somatic mutations of the CD95 gene in human B cells as a side-effect of the germinal center reaction  
JOURNAL Unpublished  
AUTHORS Muschen, M., Re, D., Brauminger, A., Wolf, J., Hansmann, M.L., Diehl, V., Kuper, R. and Rajewsky, K.  
TITLE Somatic mutations of the CD95 gene in Hodgkin- and Reed-Sternberg cells  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 702)  
AUTHORS Muschen, M.  
TITLE Direct Submission  
JOURNAL Submitted (19-SEP-2000) Muschen M., Department of Immunology, Institute for Genetics, LFI B4 R705, Joseph-Stelzmann-Str. 9, 50931 Koeln, GERMANY  
COMMENT Related sequences: D31968 X89101 AJ279012 AJ279013.  
FEATURES  
source location/Qualifiers  
1..702  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/cell\_type="Reed-Sternberg cells"  
/note="Hodgkin's disease tissue"  
1..702  
/gene="CD95"  
1..313  
/gene="CD95"  
/standard\_name="Apo-1 Fas"  
/number=1  
1..281  
/gene="CD95"  
/standard\_name="Apo-1 Fas"  
282..313  
/gene="CD95"  
/standard\_name="Apo-1 Fas"

BASE COUNT 107 a 198 c 264 g 133 t

ORIGIN

Query Match 21.7%; Score 697.2; DB 9; Length 702;  
Best local similarity 99.6%; Pred. No. 7,4e-155; Indels 0; Gaps 0;  
Matches 699; Conservative 0; Mismatches 3;

1946 GAGGCTTCTTCCATCTCTCTGACCAACGAGGCTTTGCTGAGCTGTCTGATCTCG 2005  
1 GAGGCTTCTTCCATCTCTCTGACCAACGAGGCTTTGCTGAGCTGTCTGATCTCG 60

QY 2006 CGCAAGATGACACACAGGTGTTCAAGACGCTTCTGGGAGTGAAGGAGCGGTTACG 2065

Db 61 CGCAAGATGACACACAGGTGTTCAAGACGCTTCTGGGAGTGAAGGAGCGGTTACG 120

QY 2066 AGTGAATTGCTGAGAGCTTCAAGGAGGCGGCGACCTGACAGAACACCTGAGGCGAGCC 2125

Db 121 AGTGAATTGCTGAGAGCTTCAAGGAGGCGGCGACCTGACAGAACACCTGAGGCGAGCC 180

QY 2126 CTGGGCTCCAGGCGGAGCTGCTTCTTCCCGGGGTTGTGACCCGCTCAGTACGGA 2185

Db 181 CTGGGCTCCAGGCGGAGCTGCTTCTTCCCGGGGTTGTGACCCGCTCAGTACGGA 240

QY 2186 GTTGGGAAAGCTTTTCACTTGGAGGATGCTCAACACATGCTGGGATCTGAGCC 2245

Db 241 GTTGGGAAAGCTTTTCACTTGGAGGATGCTCAACACATGCTGGGATCTGAGCC 300

QY 2246 TCTTACCTCTGATGATCCCTCTCTCTGCGGCGGTGAGGCTTACCCGCTTATGCTCCGG 2305

Db 301 TCTTACCTCTGATGATCCCTCTCTCTGCGGCGGTGAGGCTTACCCGCTTATGCTCCGG 360

QY 2306 GATAGGCAAAAGTGGGCGGCGGCGGAGCGGTGCGGAGATGCGGCGGCGAGCGC 2365

Db 361 GATAGGCAAAAGTGGGCGGCGGCGGAGCGGTGCGGAGATGCGGCGGCGAGCGC 420

QY 2366 GGGGACCTTGGAGCGCGGCTGCTGGGAGGCGGTGAGAGCTGCTCCGGGGCTGT 2425

Db 421 GGGGACCTTGGAGCGCGGCTGCTGGGAGGCGGTGAGAGCTGCTCCGGGGCTGT 480

QY 2426 TAGGACCTTCCCTCAGCGCCGGGTGCTCAGAACGATGAGAGACTTTTCTTGGGCT 2485

Db 481 TAGGACCTTCCCTCAGCGCCGGGTGCTCAGAACGATGAGAGACTTTTCTTGGGCT 540

QY 2486 TGATGCGAAGTGTGATCCCGCTGGGCGAGCGGGGAGCTCCGGGCTCTCGAGAGCA 2545

Db 541 TGATGCGAAGTGTGATCCCGCTGGGCGAGCGGGGAGCTCCGGGCTCTCGAGAGCA 600

QY 2546 CTGCGCTCCACGTTGATGAGTGGCGGTGGGGGCGGACAGAAATTGAAGCGGAGCTCGGA 2605

Db 601 CTGCGCTCCACGTTGATGAGTGGCGGTGGGGGCGGACAGAAATTGAAGCGGAGCTCGGA 660

QY 2606 AGCTTTAGGCTGCTGAGAGGGGAGCCCGGTTGAGAGAGGA 2647

Db 661 AGCTTTAGGCTGCTGAGAGGGGAGCCCGGTTGAGAGAGGA 702

RESULT 15  
HSA279013 702 bp DNA linear PRI 02-APR-2001  
LOCUS Homo sapiens partial CD95 gene for CD95 antigen (Apo-1 Fas), exon  
DEFINITION 1, polymorphic allele (+337).  
ACCESSION AJ279013 GI:13539242  
VERSION AJ279013.1  
KEYWORDS Apo-1 Fas; CD95 antigen; CD95 gene.



Sun Aug 3 09:03:38 2003

US-09-834-291-1.rst

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:13:22 ; Search time 10997.1 Seconds  
7998.748 Million cell updates/sec

Title: US-09-834-291-1  
Perfect score: 3212  
Sequence: 1 tgcagactctcaggaatgtaatg.....ttgcagatgactaatcaag 3212

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database:
- 1: em\_estdb:\*
  - 2: em\_estdb:\*
  - 3: em\_estdb:\*
  - 4: em\_estdb:\*
  - 5: em\_estdb:\*
  - 6: em\_estdb:\*
  - 7: em\_estdb:\*
  - 8: em\_estdb:\*
  - 9: gb\_est1:\*
  - 10: gb\_est2:\*
  - 11: gb\_est3:\*
  - 12: gb\_est4:\*
  - 13: gb\_est5:\*
  - 14: gb\_est6:\*
  - 15: em\_estdb:\*
  - 16: em\_estdb:\*
  - 17: em\_gss\_hum:\*
  - 18: em\_gss\_hum:\*
  - 19: em\_gss\_hum:\*
  - 20: em\_gss\_vtl:\*
  - 21: em\_gss\_fun:\*
  - 22: em\_gss\_mus:\*
  - 23: em\_gss\_mus:\*
  - 24: em\_gss\_pro:\*
  - 25: em\_gss\_pro:\*
  - 26: em\_gss\_pro:\*
  - 27: em\_gss\_vtl:\*
  - 28: gb\_gss1:\*
  - 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query Match Length	ID	Description
No.				
1	710	793	12	BT763679
2	453.4	14.1	9	AA704610
3	398	12.4	13	BQ072834
4	394.6	12.3	9	AL540709

5	386.4	12.0	506	10	BG659530	BK472931 DKFZP686B
6	381.4	11.9	506	13	BF569266	BF569266 602185634
7	357.8	11.1	599	10	BQ924045	BQ924045 AGENCOURT
8	346.2	10.8	1499	13	CB157277	CB157277 K-EST0216
9	343.6	10.7	524	14	CB157277	BG121070 602352793
10	340.6	10.6	1140	10	BG121070	AV695647 AV695647
11	328.2	10.2	617	9	AV695647	CB124167 K-EST0172
12	306.6	9.5	394	14	CB124167	AA058563 ZF55D04.8
13	291.4	9.1	603	9	AA058563	AA058563 ZF55D04.8
14	287.4	8.9	577	9	AA056275	AA047220 ZF49e11.8
15	280.4	8.7	547	9	AA047220	AA011028 ZF49d02.8
16	279.4	8.7	442	9	AA011028	AA057418 ZF59h03.8
17	278.4	8.7	447	9	AA057418	AA020992 ZF65F07.8
18	276.4	8.6	467	9	AA020992	AL711161 DKFZP686G
19	272	8.5	451	9	AL711161	BO881176 AGENCOURT
20	266.4	8.3	933	13	BO881176	H84298 YB95e09.81
21	263	8.2	376	14	H84298	H86126 YB94G08.81
22	257	8.0	398	14	H86126	AW239285 XDB3802.Y
23	256.4	8.0	429	9	AW239285	CB262638 1005143.H
24	254.4	7.9	261	14	CB262638	AA014441 ZF50a08.8
25	253	7.9	297	9	AA018441	AA014441 ZF50a08.8
26	234.8	7.3	929	13	H86544	H86544 YF04101.81
27	223.6	7.0	1201	9	BUS53711	BUS53711 AGENCOURT
28	215.4	6.7	1339	10	AL542093	AL542093 AL542093
29	214.4	6.7	1339	10	AV692383	AV692383 AV692383
30	210.4	6.6	300	9	BG288747	BG288747 602385566
31	210.4	6.6	300	9	AU100132	AU100132 AU100132
32	202.4	6.3	1052	12	BM455788	BM455788 AGENCOURT
33	200.6	6.2	741	14	AL707941	AL707941 DKFZP686D
34	200.2	6.2	741	14	CS521975	CS521975 AGENCOURT
35	199.8	6.2	775	13	B0934815	B0934815 AGENCOURT
36	192.8	6.0	800	12	CB062973	CB062973 4013286.B
37	191	5.9	800	12	CB062973	BA163384 603204413
38	190	5.9	800	12	CB062973	BA163384 603204413
39	185.2	5.8	696	9	AV715411	AV715411 YB722910.81
40	181.8	5.7	828	12	AV715411	AV715411 AV715411
41	181.4	5.6	600	14	AV715411	AV715411 AV715411
42	176.4	5.5	1152	12	AV715411	AV715411 AV715411
43	135	4.2	577	12	AV715411	AV715411 AV715411
44	131	4.1	658	13	AV715411	AV715411 AV715411
45	124.4	3.9	291	9	AV715411	AV715411 AV715411

ALIGNMENTS

RESULT 1  
BT763679/c  
LOCUS 603049567F1 NIH\_MGC\_116 Homo sapiens CDNA clone IMAGE:5189752 5',  
DEFINITION mRNA sequence.  
ACCESSION BT763679.1 GI:15755257  
VERSION BT763679.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 793)  
NITH-MGC http://imgc.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: ggaaps-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: L14M11474 row: e column: 17  
High quality sequence step: 786.







Sun Aug 3 09:03:38 2003

us-09-834-291-1.fst

TITLE EST (Blocker, H., Boecker, M., Mewes, H.W., Well, B., Amdt, C., et al.)  
JOURNAL Unpublished  
COMMENT Contact: Blocker H

MIPS Ingolstaedter Landstr. 1, D-85764 Neuberg, Germany  
This is the 5' sequence of the clone insert  
This is from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de  
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de  
sequenced by GFP (National Research Centre for Biotechnology of the  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
51 sequence also available.  
This clone (DKFZ686809151) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES  
source 1. 506  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZ686809151"  
/tissue\_type="human skeletal muscle"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_id="686 (synonym: h1cc3)"  
/note="Vector: pTriplex2; Site\_1: SfilA; Site\_2: SfilB;  
cDNA-collection"

BASE COUNT 107 a 116 c 120 g 163 t  
ORIGIN

Query Match 11.9%; Score 381.4; DB 13; Length 506;  
Best Local Similarity 98.5%; Pred. No. 2,4e-55; Indels 0; Gaps 0;  
Matches 385; Conservative 0; Mismatches 6

2822 TTAAGCCGAGGCGCTCGAAGAACGACCTTTCTTCTGAAAAGTTATGAGGGC 2881  
1 TTAAGCCGAGGCGCTCGAAGAACGACCTTTCTTCTGAAAAGTTATGAGGGC 2941  
2882 TGAATGAGCTTCTGAGAGCTTTGTTACCGTTTATTTCTCAACAGAAAAGAACTGC 120  
61 TGAATGAGCTTCTGAGAGCTTTGTTACCGTTTATTTATGACACAGAAAAGAACTGC 3001  
2942 CTGGTCCCTCCGAGGATTCCTCTTTAAGACTGTAAGTCGCTGAGAGGTTTCA 180  
121 CTGGTCCCTCCGAGGATTCCTCTTTAAGACTGTAAGTCGCTGAGAGGTTTCA 3061  
3002 TTTGTTGTTTCTGAGCTTTCTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 240  
181 TTTGTTGTTTCTGAGCTTTCTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3121  
3062 ATGATGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
241 ATGATGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3181  
3122 CCGGCGCCCTATTTATGCGCAAGAACTTGAAGAGCTGTTGAAAAGTCCCTGCTCA 360  
301 CCGGCGCCCTATTTATGCGCAAGAACTTGAAGAGCTGTTGAAAAGTCCCTGCTCA 3122  
3182 GAATGCGAGCTTGCAGATGCTGTAATCAAG 3212  
361 GAATGCGAGCTTGCAGATGCTGTAATCAAG 391

RESULT 7  
BF569266 599 bp mRNA linear EST 12-DEC-2000  
LOCUS 602185634F1 NIH\_MGC\_45 Homo sapiens cDNA clone IMAGE:4309953 5'  
DEFINITION mRNA sequence.  
ACCESSION BF569266 GI:11642749  
VERSION BF569266.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 599)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: gscaps-remail.nih.gov  
Tissue Procurement: Lindehan  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
plate: LNCM1184 row: c column: 10  
High quality sequence stop: 595.  
Location/Qualifiers

FEATURES  
source 1. 599  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4309953"  
/tissue\_type="renal carcinoma (ascites)"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NIH\_MGC\_45"  
/note="Organ: kidney; Vector: pONT7; Site\_1: XhoI; Site\_2:  
EcoRI; XhoI sites using the following 5' adaptor:  
into EcoRI/XhoI library constructed by Ling Hong at the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library"

BASE COUNT 118 a 210 c 172 g 99 t  
ORIGIN

Query Match 11.1%; Score 357.8; DB 10; Length 599;  
Best Local Similarity 99.4%; Pred. No. 2,6e-51; Indels 0; Gaps 0;  
Matches 359; Conservative 0; Mismatches 2

2289 CCGCTCTTACTCCCGGGATGAGCAAGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 302  
361 CCGCTCTTACTCCCGGGATGAGCAAGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2408  
2349 GCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 242  
301 GCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2468  
2409 TGGCTCCCGGGGCTGTTAGACCTTCCCTCAGGCGCGGCTGCTCAAGACGCTGAGGAC 182  
241 TGGCTCCCGGGGCTGTTAGACCTTCCCTCAGGCGCGGCTGCTCAAGACGCTGAGGAC 2528  
2469 TTTGTTTCTTGGGCTTGTATGCGAAGTGTGCTGCGCTGCGGCGGCGGCGGCGGCGGCGG 122  
181 TTTGTTTCTTGGGCTTGTATGCGAAGTGTGCTGCGCTGCGGCGGCGGCGGCGGCGGCGG 2588  
2529 GCGCTCTCGAGACCACTGCTCCACGTTGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGG 62  
121 GCGCTCTCGAGACCACTGCTCCACGTTGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGG 2648  
2589 GAAGCGGAAGCTTGGAGAGCTTTAGGCTGCTGAGAGGAGAGAGAGAGAGAGAGAGAG 2  
61 GAAGCGGAAGCTTGGAGAGCTTTAGGCTGCTGAGAGGAGAGAGAGAGAGAGAGAGAGAG 2

RESULT 8  
BQ924045 1499 bp mRNA linear EST 20-AUG-2002  
LOCUS BQ924045





Query Match 10.7%; Score 343.6; DB 14; Length 524;  
 Best Local Similarity 98.9%; Pred. No. 7,2e-49;  
 Matches 346; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

2289 CCGCTTATGTCCTCCGGGATAGCAAAAGTGGGGCGGGCGCGGAGCCGTGGGGATTGCG 2348  
 356 CCGCTTATGTCCTCCGGGATAGCAAAAGTGGGGCGGGCGCGGAGCCGTGGGGATTGCG 297  
 2349 GCGGACGCGGCGCACTGGGCACTGGGAGCGGGCGCTGCTGGGAGCGCTTGAAGAC 2408  
 296 GCGGACGCGGCGCACTGGGCACTGGGAGCGGGCGCTGCTGGGAGCGCTTGAAGAC 237  
 2409 TGGCTCCCGGGGGCTGTAGAACCTTCCCTCAGGCGGGGTGCTCAAGAACATGAGAGAC 2468  
 236 TGGCTCCCGGGGGCTGTAGAACCTTCCCTCAGGCGGGGTGCTCAAGAACATGAGAGAC 177  
 2469 TTGCTTTCTTGGGCTTGAATGCGAAGTGCATCCCGCTGGGCAAGCGGGCAAGCTCCG 2528  
 176 TTGCTTTCTTGGGCTTGAATGCGAAGTGCATCCCGCTGGGCAAGCGGGCAAGCTCCG 117  
 2529 GCGCTCCCTCGAGACCACTGGCTCCCACTTGAAGTGGCGCTGGGGGGCGGACAGAAAT 2588  
 116 GCGCTCCCTCGAGACCACTGGCTCCCACTTGAAGTGGCGCTGGGGGGCGGACAGAAAT 57  
 2589 GAAGCGGAAGTCTGGGAAGCTTTAGGGTGGCTGAGAGGGGAGACCCCGGTTG 2638  
 56 GAAGCGGAAGTCTGGGAAGCTTTAGGGTGGCTGAGAGGGGAGACCCCGGTTG 7

RESULT 10 1140 bp mRNA linear EST 30-JAN-2001  
 BG121070  
 LOCUS 602352793f1 NIH\_MGC\_90 Homo sapiens cDNA clone IMAGE:4451039 5',  
 DEFINITION mRNA sequence.

ACCESSION BG121070  
 VERSION BG121070.1 GI:12614579  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens (human); Chordata; Craniata; Vertebrata; Euteleostomi;  
 Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1140)  
 Mammalia; Euteleostomi; Primates; Catarrhini; Homnidae; Homo.  
 NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov/  
 Plate: L1AM10237 row: 1 column: 24  
 High quality sequence start: 9  
 High quality sequence stop: 727.  
 Location/Qualifiers

FEATURES  
 source

1. .1140  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4451039"  
 /issue\_type="adenocarcinoma, cell line"  
 /lab\_host="DHIOB (phage-resistant)"  
 /clone\_id="NIH\_MGC\_90"  
 /note="Organ: liver; Vector: pCMV-Sport6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.7 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH MGC Library."

BASE COUNT

289 a 342 c 331 g 178 t

## ORIGIN

Query Match 10.6%; Score 340.6; DB 10; Length 1140;  
 Best Local Similarity 98.4%; Pred. No. 2,2e-48;  
 Matches 365; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

2289 CCGCTTATGTCCTCCGGGATAGCAAAAGTGGGGCGGGCGCGGAGCCGTGGGGATTGCG 2348  
 375 CCGCTTATGTCCTCCGGGATAGCAAAAGTGGGGCGGGCGCGGAGCCGTGGGGATTGCG 316  
 2349 GCGGACGCGGCGCACTGGGCACTGGGAGCGGGCGCTGCTGGGAGCGCTTGAAGAC 2408  
 315 GCGGACGCGGCGCACTGGGCACTGGGAGCGGGCGCTGCTGGGAGCGCTTGAAGAC 256  
 2409 TGGCTCCCGGGGGCTGTAGAACCTTCCCTCAGGCGGGGTGCTCAAGAACATGAGAGAC 2468  
 255 TGGCTCCCGGGGGCTGTAGAACCTTCCCTCAGGCGGGGTGCTCAAGAACATGAGAGAC 196  
 2469 TTGCTTTCTTGGGCTTGAATGCGAAGTGCATCCCGCTGGGCAAGCGGGCAAGCTCC 2527  
 195 TTGCTTTCTTGGGCTTGAATGCGAAGTGCATCCCGCTGGGCAAGCGGGCAAGCTCC 136  
 2528 GCGCTCCCTCGAGACCACTGGCTCCCACTTGAAGTGGCGCTGGGGGGCGGACAGAAAT 2587  
 135 GCGCTCCCTCGAGACCACTGGCTCCCACTTGAAGTGGCGCTGGGGGGCGGACAGAAAT 76  
 2588 TGAAGCGGAAGTCTGGGAAGCTTTAGGGTGGCTGAGAGGGGAGACCCCGTTGAGAGAG 2646  
 75 TGAAGCGGAAGTCTGGGAAGCTTTAGGGTGGCTGAGAGGGGAGACCCCGTTGAGAGAG 16  
 2647 AGCGGAAGCTCC 2657  
 15 AGCGGAAGCTCC 5

RESULT 11 617 bp mRNA linear EST 16-JAN-2002  
 AV695647  
 LOCUS AV695647 GKC Homo sapiens cDNA clone GKCPCD11 5', mRNA sequence.

DEFINITION AV695647  
 ACCESSION AV695647  
 VERSION AV695647.1 GI:10297510  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens (human); Chordata; Craniata; Vertebrata; Euteleostomi;  
 Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 617)  
 Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X.,  
 Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W.,  
 Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X.,  
 Hu, G., Gu, J., Chen, Z., and Han, Z.  
 TITLE Insight into hepatocellular carcinogenesis at transcriptome level  
 by comparing gene expression profiles of hepatocellular carcinoma  
 with those of corresponding noncancerous liver

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)  
 MEDLINE 21625106  
 PUBMED 11752456  
 COMMENT Contact: Zeguang Han  
 Chinese National Human Genome Center at Shanghai  
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
 201203, P. R. China  
 Tel: 86-21-50801919 (ex.45)  
 Fax: 86-21-50801922  
 Email: hanzg@chgc.sh.cn  
 This clone is available at CHGC in Shanghai.

## FEATURES

source  
 1. .617  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="GKCPCD11"  
 /issue\_type="hepatocellular carcinoma"  
 /dev\_stage="Adult"

```

/lab host="SOLR"
/clone.lib="GKC"
/Note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT      122 a      210 c      179 g      104 t      2 others
ORIGIN

```

```

Query Match      10.2%; Score 328.2; DB 9; Length 617;
Best Local Similarity 98.0%; Pred. No. 3.1e-46;
Matches 344; Conservative 0; Mismatches 3; Indels 4; Gaps 1;

```

```

QY 2289 CCGCTTACTTCCCGGGATGAGCAAAAGTGGGCGGGCGGGCGGGCGGGATTTGCG 2348
DB 347 CCGCTTACTTCCCGGGATGAGCAAAAGTGGGCGGGCGGGCGGGATTTGCG 288
QY 2349 GCGGACGCGGCGGACCGGCGGACCTTGGAGCCGGGCTGCTGGAGAGCGTTGAGAGC 2408
DB 287 GCGGACGCGGCGGACCGGCGGACCTTGGAGCCGGGCTGCTGGAGAGCGTTGAGAGC 228
QY 2409 TGGCTCCCGGGGCTTTAGAACCTTCCCTCAGCGCGGGCTGCTGAGAACATGAGAGC 2468
DB 227 TGGCTCCCGGGGCTTTAGAACCTTCCCTCAGCGCGGGCTGCTGAGAACATGAGAGC 168
QY 2469 TTGCTTTTCTTGGGCTTGAATGAGAGTGTGATCCCGCTGGGCGAGCGGCGAGCTCCG 2528
DB 167 TTGC-----CTTGGGCTTGAATGAGAGTGTGATCCCGCTGGGCGAGCGGCGAGCTCCG 112
QY 2529 GCGGCTCCCGGAGACCACTGCGCTTCACGTTGAGGTGGGCGGGCGGAGCAAGATT 2588
DB 111 GCGGCTCCCGGAGACCACTGCGCTTCACGTTGAGGTGGGCGGGCGGAGCAAGATT 52
QY 2589 GAAGCGGAAGTCTGGAAGCTTTAGGGTCTGAGAGGGGAGACCCCGGTTGG 2639
DB 51 GAAGCGGAAGTCTGGAAGCTTTAGGGTCTGAGAGGGGAGACCCCGGTTGG 1

```

```

RESULT 12      394 bp      mRNA      linear      EST 29-JAN-2003
LOCUS      CB124167/c
DEFINITION      K-EST0172829 LSHLK1s1 Homo sapiens cDNA clone LSHLK1s1-11-C12 5',
VERSION      CB124167
KEYWORDS      EST.
SOURCE      CB124167.1 GI:28084072
ORGANISM      Homo sapiens (human)

```

```

REFERENCE      1 (bases 1 to 394)
AUTHORS      Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
              Kim,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
              21C Frontier Korean EST Project 2001

```

```

TITLE      Unpublished
JOURNAL      Contact: Kim YS
COMMENT      Genome Research Center
              Korea Research Institute of Bioscience & Biotechnology
              52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
              Tel: +82-42-860-4470
              Fax: +82-42-860-4409
              Email: yongsung@mail.kribb.re.kr
              Plate: 11 row: C column: 12
              High quality sequence stop: 394.

```

```

FEATURES
Source
1..394
Location/Qualifiers

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="LSHLK1s1-11-C12"
/sex="W"
/cell_line="HLK-1"
/lab_host="Top10F"
/clone.lib="LSHLK1s1"

```

```

/Note="Organ: Liver; Vector: pCNS-D2; Site_1: EcoRI;
Site_2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then dephosphorylated with
with tabacco acid pyrophosphatase (TAP) and then dephosphorylated with
TAP. The dephosphorylated RNA was ligated with DNA-RNA linker including
EcoRI site by treatment of T4 RNA ligase and the first
strand cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library. After analyzing and
sequencing about 2,000 - 3,000 colonies in original cDNA
library, the abundant cDNAs were selected and amplified by
PCR reaction using vector region primer including T7
promotor as 5' primer and N(drr)14 as 3' primer. The PCR
products were used as template for synthesis of
biotinylated single stranded RNA by in vitro transcription
reaction. The synthesized RNA probes were hybridized with
antisense single stranded cDNAs prepared from original
library and incubated with avidin-gel. After removing
DNA-RNA hybrids by centrifuge, the substracted cDNA
libraries were constructed by transformation of the
remaining DNA into competent cells E. coli Top10F with
electroporation method."

```

```

BASE COUNT      75 a      149 c      109 g      61 t
ORIGIN

```

```

Query Match      9.5%; Score 306.6; DB 14; Length 394;
Best Local Similarity 98.7%; Pred. No. 1.1e-42;
Matches 309; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 2289 CCGCTTACTTCCCGGGATGAGCAAAAGTGGGCGGGCGGGCGGGATTTGCG 2348
DB 315 CCGCTTACTTCCCGGGATGAGCAAAAGTGGGCGGGCGGGCGGGATTTGCG 256
QY 2349 GCGGACGCGGCGGACCGGCGGACCTTGGAGCCGGGCTGCTGGAGAGCGTTGAGAGC 2408
DB 255 GCGGACGCGGCGGACCGGCGGACCTTGGAGCCGGGCTGCTGGAGAGCGTTGAGAGC 196
QY 2409 TGGCTCCCGGGGCTTTAGAACCTTCCCTCAGCGCGGGCTGCTGAGAACATGAGAGC 2468
DB 195 TGGCTCCCGGGGCTTTAGAACCTTCCCTCAGCGCGGGCTGCTGAGAACATGAGAGC 136
QY 2469 TTGCTTTTCTTGGGCTTGAATGAGAGTGTGATCCCGCTGGGCGAGCGGCGAGCTCCG 2528
DB 135 TTGCTTTTCTTGGGCTTGAATGAGAGTGTGATCCCGCTGGGCGAGCGGCGAGCTCCG 76
QY 2529 GCGCTTCTTGGAGACCACTGCGCTTCACGTTGAGGTGGGCGGGCGGAGCAAGATT 2588
DB 75 GCGCTTCTTGGAGACCACTGCGCTTCACGTTGAGGTGGGCGGGCGGAGCAAGATT 16
QY 2589 GAAGCGGAAGTCT 2601
DB 15 GAAGCGGAAGTCT 3

```

```

RESULT 13      603 bp      mRNA      linear      EST 02-FEB-1997
LOCUS      AA058563
DEFINITION      z155b04.g1 Soares retina N2b4HR Homo sapiens cDNA clone
IMAGE:380815 3', mRNA sequence.
VERSION      AA058563
KEYWORDS      EST.
SOURCE      AA058563.1 GI:1551370
ORGANISM      Homo sapiens (human)

```

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

REFERENCE 1 (bases 1 to 603)  
 AUTHORS Hallier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,  
 Chissee, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins,  
 'M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore,  
 B., Morris, M., Parsons, J., Frange, C., Rifkin, L., Rohlfing, T.,  
 Schellenberg, K., Soares, M.B., Tan, F., Thierly-Meg, J., Trevaekis, E.,  
 Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.  
 Generation and analysis of 280,000 human expressed sequence tags  
 Genome Res. 6 (9), 807-828 (1996)

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: east@wustl.edu

High quality sequence stop: 467.  
 Insert Length: 921 Std Error: 0.00  
 Seq primer: -40M13 fwd. from Amersham  
 Location/Qualifiers

FEATURES  
 source

1.603  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:1289072"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:380815"  
 /sex="male"  
 /tissue\_type="retina"  
 /dev\_stage="55 year old"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares retina N2B4R"  
 /note="Organ: eye; Vector: pT73D (Pharmacia) with a  
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5'  
 TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTCTTTTCTTTT 3']  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pT73 vector  
 (Pharmacia). The retinas were obtained from a 55 year old  
 Caucasian and total cellular poly(A)+ RNA was extracted 6  
 hrs after their removal. The retina RNA was kindly  
 provided by Roderick R. McInnes M.D. Ph.D. from the  
 University of Toronto. Library constructed by Bento  
 Soares and M.Fatima Bonaldo."

BASE COUNT 145 a 142 c 140 g 175 t 1 others

Query Match 9.1%; Score 291.4; DB 9; Length 603;  
 Best Local Similarity 99.7%; Pred. No. 6,4e-40;  
 Matches 292; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2920 GTGACACAGAAAGAAAGAAAGCTGCTGCTCCCTCCGGAATTCCTTTAAGCTGTA 2979  
 1 GTGACACAGAAAGAAAGAAAGCTGCTGCTCCCTCCGGAATTCCTTTAAGCTGTA 60  
 2980 AGTGGCTGCTGAGTGTTCATTTGTTTTCGCTCCCTCTCTTTCTTTTGG 3039  
 61 AGTGGCTGCTGAGTGTTCATTTGTTTTCGCTCCCTCTCTTTCTTTTGG 120  
 3040 CCCTTTCTAGCTTGCATCCCATGATGATTTGCTGCTGCTCCGCGGTTGGTGG 3099  
 121 CCCTTTCTAGCTTGCATCCCATGATGATTTGCTGCTGCTCCGCGGTTGGTGG 180  
 3100 TACTGCTTCCACCGGACAGAAACCGGCGCTATTTATTTGCAAGAACTTGAAGACCT 3159  
 181 TACTGCTTCCACCGGACAGAAACCGGCGCTATTTATTTGCAAGAACTTGAAGACCT 240  
 3160 GTTTTGAAGTCCCTGCTGCTGCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3212  
 241 GTTTTGAAGTCCCTGCTGCTGCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 293

RESULT 14  
 AA056275 577 bp mRNA linear EST 17-SEP-1996  
 LOCUS 2153903.81 Soares retina N2B4R Homo sapiens cDNA clone  
 DEFINITION IMAGE:380692.3, mRNA sequence.

ACCESSION  
 AA056275  
 VERSION  
 AA056275.1 GI:1548679  
 EST.  
 Homo sapiens (human)  
 SOURCE  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 577)  
 AUTHORS Hallier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman,  
 'M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,  
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaekis, E., Waterston,  
 R., Williamson, A., Wohlmann, P. and Wilson, R.  
 The WashU-Merck EST Project  
 Unpublished  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: east@wustl.edu

High quality sequence stop: 178.  
 Insert Length: 921 Std Error: 0.00  
 Seq primer: -40M13 fwd. from Amersham  
 Location/Qualifiers

FEATURES  
 source

1.577  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:1288949"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:380692"  
 /sex="male"  
 /tissue\_type="retina"  
 /dev\_stage="55 year old"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares retina N2B4R"  
 /note="Organ: eye; Vector: pT73D (Pharmacia) with a  
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5'  
 TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTCTTTTCTTTT 3']  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pT73 vector  
 (Pharmacia). The retinas were obtained from a 55 year old  
 Caucasian and total cellular poly(A)+ RNA was extracted 6  
 hrs after their removal. The retina RNA was kindly  
 provided by Roderick R. McInnes M.D. Ph.D. from the  
 University of Toronto. Library constructed by Bento  
 Soares and M.Fatima Bonaldo."

BASE COUNT 133 a 135 c 135 g 163 t 11 others

Query Match 8.9%; Score 287.4; DB 9; Length 577;  
 Best Local Similarity 98.3%; Pred. No. 3.1e-39;  
 Matches 288; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

2920 GTGACACAGAAAGAAAGAAAGCTGCTGCTCCCTCCGGAATTCCTTTAAGCTGTA 2979  
 1 GTGACACAGAAAGAAAGAAAGCTGCTGCTCCCTCCGGAATTCCTTTAAGCTGTA 60  
 2980 AGTGGCTGCTGAGTGTTCATTTGTTTTCGCTCCCTCTCTTTCTTTTGG 3039  
 61 AGTGGCTGCTGAGTGTTCATTTGTTTTCGCTCCCTCTCTTTCTTTTGG 120  
 3040 CCCTTTCTAGCTTGCATCCCATGATGATTTGCTGCTGCTCCGCGGTTGGTGG 3099

Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.jhmi.gov](mailto:info@image.jhmi.gov)) for further information.  
Seq length: 893 Std Error: 0.00  
Seq primer: -40M13 fwd. from Amerstham  
High quality sequence stop: 471.  
Location/Qualifiers  
1. 547  
Source

131 a	128 c	125 g	158 t	5 others
-------	-------	-------	-------	----------

Sun Aug 3 09:03:34 2003

US-09-834-291-1.rng

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Inc.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:32:17 ; Search time 1075.81 Seconds  
(without alignments)  
8059.612 Million cell updates/secTitle: US-09-834-291-1  
Perfect score: 3212  
Sequence: 1 tggagacttcaggaatg.....ttgcagatgctatcaag 3212Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0Searched: 2552756 seqs, 1349719017 residues  
Total number of hits satisfying chosen parameters: 5105512Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

1: N Geneseq.19jun03:\*

2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*

3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*

4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*

5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*

6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*

7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*

8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*

9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*

10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*

11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*

12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*

13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*

14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*

15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*

16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*

17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*

18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*

19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*

20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*

21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*

22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*

23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001.DAT:\*

24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2143.4	66.7	2165	ABN79677	Sequence #1 used to
2	1536.8	47.8	1608	ABT34162	Fas promoter regio
3	961	29.9	1608	ABJ34419	Human immune syste
4	957.4	29.8	1608	ABJ34418	Human immune syste
5	287.4	8.9	21	AAJ98177	Human colon cancer
6	267.8	8.2	266	AAJ98177	Human CD5 recepto
7	252	7.8	2551	AAJ98177	DNA encoding a hum
8	252	7.8	2551	ABV94152	Breast carcinoma t

9	252	7.8	2551	24	ABN79588	Human Fas locus HS
10	252	7.8	2551	25	ABX10416	DNA encoding prote
11	226	7.0	1024	25	AB283782	Toxicologically re
12	226	7.0	2471	16	AAQ93879	Fas-delta-TM cDNA.
13	226	7.0	2471	20	AAJ24878	Soluble Fas recept
14	226	7.0	2534	13	AAQ29959	Human cell surface
15	226	7.0	2534	15	AAQ95297	Plasmod pf58 contg
16	226	7.0	2534	17	AAJ16303	hFas coding sequen
17	226	7.0	2534	17	AAJ29993	Fas cDNA. Mammali
18	224.4	7.0	2534	18	AAV32993	Human Fas antigen
19	224.4	7.0	2534	18	AAV32993	Human Fas antigen
20	60.4	1.9	260209	24	ABJ56564	Human lung tumour
21	60	1.9	2435	22	AAJ68854	Human lung cancer
22	60	1.9	2435	25	AAJ1094	Human lung cancer
23	60	1.9	2435	25	AAJ1094	Human lung cancer
24	60	1.9	2821	24	ABK39630	Human lung cancer
25	60	1.9	2821	25	AAJ11959	Human lung cancer
26	60	1.9	2821	25	AAJ11959	Human lung cancer
27	59.8	1.9	35100	19	AAJ20441	Human c-fms oncoge
28	59.8	1.9	35100	19	AAJ20441	Human c-fms oncoge
29	59.8	1.9	35100	19	AAJ20441	Human c-fms oncoge
30	58.6	1.8	8888	20	AAJ2665	Human cDNA differe
31	58.4	1.8	77425	24	ABK3502	Human cDNA differe
32	58.4	1.8	9947	24	ABK3502	Human cDNA differe
33	57.2	1.8	43599	24	ABK3502	Human cDNA differe
34	57	1.8	857	17	AAJ4530	Human Fas soluble
35	57	1.8	920	17	AAJ4530	Human Fas soluble
36	57	1.8	920	17	AAJ4530	Human Fas soluble
37	57	1.8	920	17	AAJ4530	Human Fas soluble
38	57	1.8	975	24	ABJ63728	Human Fas soluble
39	57	1.8	975	24	ABJ63728	Human Fas soluble
40	57	1.8	1104	17	AAJ34526	Human Fas soluble
41	57	1.8	1167	17	AAJ34526	Human Fas soluble
42	56.8	1.8	62488	24	ABJ4981	Human Fas soluble
43	56.4	1.8	2502	25	ABJ76239	Human Fas soluble
44	56.2	1.7	35901	20	AAJ23892	Human Fas soluble
45	56.2	1.7	38886	20	AAJ23892	Human Fas soluble

## ALIGNMENTS

RESULT 1  
ID ABN79677 standard; DNA; 2165 BP.

AC ABN79677; (first entry)

DT 29-JUL-2002 (first entry)

DE Sequence #1 used to generate target oligonucleotides.

DE Human; immunosuppressive; antiinflammatory; hepatotropic;

KW Cytostatic; vasotropic; hepatitis; cancer; allograft rejection;

KW de; Fas.

OS Homo sapiens.

XX Key

XX CDS

XX CD5

XX CD5

XX CD5

XX CD5

US2002004490-A1.

10-JAN-2002.

09-MAR-2001; 2001US-0802669.

12-APR-1999; 99US-0290640.

18-SEP-2000; 2000US-0665615.

18-SEP-2000; 2000US-0665615.

18-SEP-2000; 2000US-0665615.

Location/Qualifiers  
1782..1813  
/tag= a  
/product= "peptide encoded by sequence  
used create target oligonucleotides"





```

Db      67 GAGTGTGTGACAAAGGCTGGCAGCCGAGGCTCTTCTCACTAGCCATCAAGTCACTGCA 126
Qy      887 AAGGTGGAACAGAGACAGGCTATCAACACTTCAAGACTGTGTGTAGTGAAGTGAAGAG 126
Db      127 AAGGTGGAACAGAGACAGGCTATCAACACTTCAAGACTGTGTGTAGTGAAGTGAAGAG 946
Qy      947 ATGCAAAACACAGAGTGTGTGAAGAGGCTCTCAAGAGGTTACCTTAACCTTAAGTGAAGG 186
Db      187 ATGCAAAACACAGAGTGTGTGAAGAGGCTCTCAAGAGGTTACCTTAACCTTAAGTGAAGG 1006
Qy      1007 CC-AACAGGCTCCAGAAAGAAATGTCACTGAGAGAAAGCTGAAAGATGAACAGTGGGC 246
Db      247 CCAGAGGCTCCAGAAAGAAATGTCACTGAGAGAAAGCTGAAAGATGAACAGTGGGC 306
Qy      1066 TAAAGAGGTTTATTAATGTGTATTAATGTGTGTATTAATGTGTGTATTAATGTGTGTATTA 1125
Db      307 TAAAGAGGTTTATTAATGTGTATTAATGTGTGTATTAATGTGTGTATTAATGTGTGTATTA 366
Qy      1126 TTGCAAGTGAAGTGCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1185
Db      367 TTGCAAGTGAAGTGCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1245
Qy      1186 TGTCCAGTCTGGAACCTGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 426
Db      427 TGTCCAGTCTGGAACCTGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 486
Qy      1246 CCTTCGTAAATTTATGCTATTAATCACTAAGCTTATCACTTATCACTTATCACTTATCACT 1305
Db      487 CCTTCGTAAATTTATGCTATTAATCACTAAGCTTATCACTTATCACTTATCACTTATCACT 546
Qy      1306 ACTTTGAACAGTGTTCACAGAGACAGAAAGAAATTAACAAGATTTTATTAAGAAATTT 1365
Db      547 ACTTTGAACAGTGTTCACAGAGACAGAAAGAAATTAACAAGATTTTATTAAGAAATTT 606
Qy      1366 GGCAGAGAAATTAATGAGTAAACAGAGACAGAAAGAAATTAACAAGATTTTATTAAGAAATTT 1425
Db      607 GGCAGAGAAATTAATGAGTAAACAGAGACAGAAAGAAATTAACAAGATTTTATTAAGAAATTT 666
Qy      1426 GGCAGAGAAATTAATGAGTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1485
Db      667 GGCAGAGAAATTAATGAGTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1485
Qy      1486 GGCAGAGAAATTAATGAGTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 726
Db      727 GGCAGAGAAATTAATGAGTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1545
Qy      1546 TCCATTCAGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 786
Db      787 TCCATTCAGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1605
Qy      1606 AATATGCCCCGCAAGTCTTCTCTGAGTACTCCAGCAATTAAGCCAGAGCTCTGTATCCCA 846
Db      847 AATATGCCCCGCAAGTCTTCTCTGAGTACTCCAGCAATTAAGCCAGAGCTCTGTATCCCA 1665
Qy      1666 GGCAGAGAAATTAATGAGTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 906
Db      907 GGCAGAGAAATTAATGAGTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1725
Qy      1726 GAACTTACAGAGAGGCTTTTAAAGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 966
Db      967 GAACTTACAGAGAGGCTTTTAAAGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1785
Qy      1786 GTGAGCAATGCAAGCACTGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1845
Db      1845 GTGAGCAATGCAAGCACTGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1845
Qy      1846 CCTTGAATCTTCTCACTGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1905
Db      1905 CCTTGAATCTTCTCACTGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1905
Qy      1906 TGAATCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1965

```

```

Db      1146 TGAATCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1205
Qy      1966 CTGACCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1205
Db      1206 CTGACCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2025
Qy      2026 GTTCAAAAGAGGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1265
Db      1266 GTTCAAAAGAGGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2085
Qy      2086 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1325
Db      1325 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2145
Qy      2146 GCTCTTCTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1385
Db      1386 GCTCTTCTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2205
Qy      2206 TGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1445
Db      1446 TGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2265
Qy      2266 CTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1505
Db      1506 CTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2325
Qy      2326 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1565
Db      1566 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1606

RESULT 3
ABL34419/c
ID ABL34419 standard; DNA; 1608 bp.
AC ABL34419;
XX
XX
XX 26-MAR-2002 (first entry)
XX
XX Human immune system associated gene SEQ ID NO: 2392.
XX
XX Human; immune system disease; cytosine methylation; antiaesthetic;
XX antileukemic; antineoplastic; antineoplastic; antineoplastic;
XX antineoplastic; antineoplastic; antineoplastic; antineoplastic;
XX antineoplastic; antineoplastic; antineoplastic; antineoplastic;
XX acute myeloid leukemia; Alzheimer's disease; arteriosclerosis;
XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
XX gene; ds.
XX
XX Homo sapiens.
XX
XX WO200200928-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP07537.
XX
XX 30-JUN-2000; 2000DE-1032529.
XX
XX 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIC-) BEIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful
XX for diagnosis and treatment of diseases associated with abnormal
XX cytosine methylation -
XX
XX Claim 1; SEQ ID NO 2392; 32pp + Sequence Listing; German.
XX

```





XX 03-JAN-2002.  
 PD 02-JUL-2001; 2001WO-EP07537.  
 XX 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX (EPIC-) EPIGENOMICS AG.  
 PI Olek A, Piepenbrock C, Berlin K,  
 XX WPI; 2002-130909/17.  
 DR Nucleic acid comprising fragment of chemically modified gene, useful  
 XX for diagnosis and treatment of diseases associated with abnormal  
 PT cytosine methylation  
 PS Claim 1; SEQ ID NO 2391; 32bp + Sequence listing; German.

The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC leukemia, Alzheimer's disease, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention.

Sequence 1608 BP; 398 A; 51 C; 423 G; 736 T; 0 other;

Query Match

Best Local Similarity 29.8%; Score 957.4; DB 24; Length 1608;  
 Matches 1215; Conservative 0; Mismatches 371; Indels 5; Gaps 3;

779 TATACCATCCCTTATCCACTTTTGTGCTTATGAGTCTCAGAGTGTGCAC 838  
 19 TATATTTTATTTTATTTTATTTTGTGCTTATGAGTCTCAGAGTGTGCAC 838  
 839 AAGCTGCAAGCCAGGAGGCTTCCATGCGCACTAAGCTCTAGAAAGGTGCAAG 898  
 79 AAGGTGTGAGCTTGTAGGCTTTTATGAGTATTAATGATTTATGAAGGTGCAAG 898  
 899 AGACAAGCTATGACCACTAAGAGTGTGTAGTGTGAGTGTGAGTGTGAGTGTGAG 138  
 139 AGATAGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 958  
 959 GGGTATGGAAGCCCTCAGAGGAGTAACTAAGTATTTAGGCTTAAATAGCTT 198  
 199 GGGTATGGAAGCCCTCAGAGGAGTAACTAAGTATTTAGGCTTAAATAGCTT 1017  
 1018 CAGAGAAATGTCACTGAGAGGAGTAACTAAGTATTTAGGCTTAAATAGCTT 258  
 259 TAGAAGAAATGTCACTGAGAGGAGTAACTAAGTATTTAGGCTTAAATAGCTT 1077  
 1078 TATTAATGTGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 318  
 319 TATTAATGTGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1137  
 1138 GTCCAGAGCTGTGAGAGTGTGAGAGTAACTAAGTATTTAGGCTTAAATAGCTT 378  
 379 GTGTAGAGTGTGTGAGAGTGTGAGAGTAACTAAGTATTTAGGCTTAAATAGCTT 1197  
 1198 AATGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 438  
 439 AATGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1257  
 1258 TTCAATGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 498  
 499 TTATGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1317  
 1318 GTTACAGAGCAAGAAATTTCAAGATTTTATTTATTTATTTATTTATTTATTTATTT 558  
 1377

559 GTTATTAGAGTACGAAAGATTTATTAAGATTTTATTTATTAAGAAATTTGTTAGAAATA 618  
 1378 ATGATGACGAGAGAGAGAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 618  
 619 ATGATGACGAGAGAGAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1437  
 1438 TGGCTTAAGTGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 678  
 679 TGGCTTAAGTGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1497  
 1498 TTCAAGAGCTTATGAGAGAGAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 728  
 729 TTCAAGAGCTTATGAGAGAGAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1557  
 1558 AAGTGTGAGAGAGAGAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 798  
 799 AAGTGTGAGAGAGAGAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1617  
 1618 AAGTGTGAGAGAGAGAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 858  
 859 AAGTGTGAGAGAGAGAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1677  
 1678 GGGCTGTAGAGAGAGAGAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 918  
 919 GGGCTGTAGAGAGAGAGAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1737  
 1738 AAGCTTATGAG 978  
 979 AAGCTTATGAG 1797  
 1798 GGCATGAG 1037  
 1038 GTTATTTATGAG 1857  
 1858 CTTATTTATGAG 1097  
 1098 TTTATTTATGAG 1917  
 1918 GGCCTTATGAG 1157  
 1158 GTTATTTATGAG 1977  
 1978 GCTTTATGAG 1217  
 1218 GTTATTTATGAG 2037  
 2038 TTTTGGAG 1277  
 1278 TTTTGGAG 2097  
 2098 TGGCAG 1337  
 1338 TGGTACGAG 2157  
 2158 GGGGAG 1397  
 1398 GGGGAG 2217  
 2218 TCAACAG 1457  
 1458 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2277  
 2278 TGAAG 1517  
 1518 TGAAG 2337  
 2338 GCGGAG 1574  
 1575 GCGGAG 1605

RESULT 5  
 AAC98177





useful for the prognosis or diagnostic of tumour, in differentiating a normal cell from a cancer cell, detecting a hormone sensitive tumour cell, differentiating a tumour with lymph nodes from a tumour without lymph nodes, differentiating antineoplastic-sensitive tumours from antineoplastic-insensitive tumours, and classifying good and poor prognosis primary breast tumours. (1) is useful for large-scale molecular characterization of breast cancer that help in prediction, prognosis and cancer treatment, and for detecting differentially expressed genes that correlated with a cancer.

Sequence 2551 BP; 826 A; 493 C; 508 G; 724 T; 0 other;

Query Match 7.8%; Score 252; DB 24; Length 2551;  
Best Local Similarity 100.0%; Pred. No. 1.2e-57;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2007 GCAGAGTGACACACAGGTGTCMAAGACGCTTCTGGGAGTGAGGAGCGGTTTACGA 2066  
1 GCAAGAGTGACACACAGGTGTCMAAGACGCTTCTGGGAGTGAGGAGCGGTTTACGA 60  
2067 GTGACTTGGCTGAGAGCTTCAAGGCGGCGCACTGGACACGACCTTGAGGCGGCGCC 2126  
61 GTGACTTGGCTGAGAGCTTCAAGGCGGCGCACTGGACACGACCTTGAGGCGGCGCC 120  
2127 TGGCTGGCCGAGCGGAGCTGCTTCTCCCGCGGGTGGTGAACCGCTCAGTACGAG 2186  
121 TGGCTGGCCGAGCGGAGCTGCTTCTCCCGCGGGTGGTGAACCGCTCAGTACGAG 2246  
2187 TTGGGGAAGCTCTTTCACCTTGAGAGATTGCTCAACACATGCTGGGAGTGGAGCCCT 2246  
Db 181 TTGGGGAAGCTCTTTCACCTTGAGAGATTGCTCAACACATGCTGGGAGTGGAGCCCT 240  
Qy 2247 CCTACTCTGT 2258  
241 CCTACTCTGT 252  
Db

RESULT 9  
ABN79588 standard; DNA; 2551 BP.

1D ABN79588 standard; DNA; 2551 BP.  
XX AC ABN79588;  
XX 29-JUL-2002 (first entry)  
XX  
XX Human Fas locus HSAF01.  
XX  
XX Human; immunosuppressive; antiinflammatory; hepatotropic;  
XX cytosolic; vasotrophic; hepatitis; cancer; allograft rejection;  
XX ds; Fas.  
XX  
XX Homo sapiens.  
XX  
XX Key location/Qualifiers  
XX CDS 221..1228  
XX FT //tag= a  
XX FT /product= "Fas HSAF01"  
XX  
XX US2002004490-A1.  
XX  
XX 10-JAN-2002.  
XX  
XX 09-MAR-2001; 2001US-0802669.  
XX  
XX 12-APR-1999; 99US-0290640.  
XX  
XX 18-SEP-2000; 2000US-0665615.  
XX  
XX (DEAN/) DEAN N M.  
XX (MARC/) MARCUSON E G.  
XX (WYATT/) WYATT J.  
XX (ZHANG/) ZHANG H.  
XX  
XX Dean NM, Marcuson EG, Wyatt J, Zhang H;  
XX  
XX

XX WP1: 2002-20486/26.  
DR P-PSDB: ABP3562.

XX Novel antisense compound targeted to nucleic acid encoding Fas, Fas  
PT ligand or Fas associated protein-1 is useful for inhibiting expression  
PT of Fas, Fas ligand, or Fas-1 in cells or tissues, and for treating  
XX hepatitis.

PS Example 2; Page 25-27; 84pp; English.

XX This invention relates to an antisense compound encoding Fas,  
XX Fas ligand, or Fas associated protein-1 (Fas-1). The inhibition of  
XX Fas mediated signalling is thought to be immunosuppressive,  
XX antiinflammatory, hepatotropic, cytostatic and vasotrophic.  
XX Antisense oligonucleotides were designed to target human Fas.  
XX Oligonucleotides were synthesized as chimeric oligonucleotides  
XX and are useful for treating an animal having an autoimmune or  
XX inflammatory disease e.g., hepatitis, cancer, a condition associated  
XX with apoptosis, allograft rejection, or ischemia reperfusion  
XX injury. Optionally, the above mentioned conditions are prevented by  
XX contacting the allograft with the antisense oligonucleotide. The  
XX oligonucleotides are used in diagnostics, therapeutics, prophylaxis  
XX and as research reagents and in kits. The oligonucleotides are also  
XX useful for research purposes. The present nucleotide sequence is  
XX related to human Fas.

Sequence 2551 BP; 826 A; 493 C; 508 G; 724 T; 0 other;

Query Match 7.8%; Score 252; DB 24; Length 2551;  
Best Local Similarity 100.0%; Pred. No. 1.2e-57;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2007 GCAGAGTGACACACAGGTGTCMAAGACGCTTCTGGGAGTGAGGAGCGGTTTACGA 2066  
1 GCAAGAGTGACACACAGGTGTCMAAGACGCTTCTGGGAGTGAGGAGCGGTTTACGA 60  
2067 GTGACTTGGCTGAGAGCTTCAAGGCGGCGCACTGGACACGACCTTGAGGCGGCGCC 2126  
61 GTGACTTGGCTGAGAGCTTCAAGGCGGCGCACTGGACACGACCTTGAGGCGGCGCC 120  
2127 TGGCTGGCCGAGCGGAGCTGCTTCTCCCGCGGGTGGTGAACCGCTCAGTACGAG 2186  
121 TGGCTGGCCGAGCGGAGCTGCTTCTCCCGCGGGTGGTGAACCGCTCAGTACGAG 2246  
2187 TTGGGGAAGCTCTTTCACCTTGAGAGATTGCTCAACACATGCTGGGAGTGGAGCCCT 2246  
Db 181 TTGGGGAAGCTCTTTCACCTTGAGAGATTGCTCAACACATGCTGGGAGTGGAGCCCT 240  
Qy 2247 CCTACTCTGT 2258  
241 CCTACTCTGT 252  
Db

RESULT 10  
ABX10416 standard; DNA; 2551 BP.

1D ABX10416 standard; DNA; 2551 BP.  
XX AC ABX10416;  
XX 28-JAN-2003 (first entry)  
XX  
XX DNA encoding protein differentially regulated in prostate cancer #85.  
XX  
XX Prostate cancer; gene expression; differential regulation;  
XX molecular marker; drug target; cancer detection; cancer diagnosis;  
XX cancer staging; cancer grading; cancer assessing; cancer monitoring;  
XX gene; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO20020281638-A2.  
XX  
XX

Query Match 7.0%; Score 226; DB 25; Length 1024;  
Best Local Similarity 100.0%; Pred. No. 8 Aa-Et

Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2033 GACGCTTCGGGGAGTGAAGGAGCGTTTACAGTGAAGCTTGAGAGCTCAGGGGCG 2092  
 1 GACGCTTCGGGGAGTGAAGGAGCGTTTACAGTGAAGCTTGAGAGCTCAGGGGCG 60

DB 2093 GGCACCTGGACGAGAACACACCTCTGAGGCGCAGCCCTGGCTGCCAGGCGAGCTGCTCTT 2152  
 61 GGCACCTGGACGAGAACACACCTCTGAGGCGCAGCCCTGGCTGCCAGGCGAGCTGCTCTT 120

QY 2153 CTCCTGGCGGGTGTGTGAGACCCGCTCAGTACGAGAGTGGAGAGCTTTCACTTCGAGG 2212  
 121 CTCCTGGCGGGTGTGTGAGACCCGCTCAGTACGAGAGTGGAGAGCTTTCACTTCGAGG 180

DB 2213 ATTGCTCAACACCATGCTGGGCACTTGACCCCTCTCACTCTGGT 2258  
 181 ATTGCTCAACACCATGCTGGGCACTTGACCCCTCTCACTCTGGT 226

QY 2213 ATTGCTCAACACCATGCTGGGCACTTGACCCCTCTCACTCTGGT 2258  
 181 ATTGCTCAACACCATGCTGGGCACTTGACCCCTCTCACTCTGGT 226

## RESULT 12

AAQ93879 ID AAQ93879 standard; cDNA; 2471 BP.

AC AAQ93879;

XX 25-MAR-2003 (updated)

DT 06-NOV-1995 (first entry)

XX Fas-delta-TM cDNA.

XX Fas-delta-TM; transmembrane deletion; apoptosis; antibody;  
 adoptive immunotherapy; transgenic animal; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 195..1139

FT sig\_peptide /tag= a

FT mat\_peptide /tag= b

FT /tag= c

XX MO9513701-A1.

XX 26-MAY-1995.

XX 15-NOV-1994; 94WO-US13173.

XX 15-NOV-1993; 93US-0152443.

XX (LXRB-) LXR BIOTECHNOLOGY INC.

XX Barr PJ, Kiefer MC, Shapiro JP;

XX WPI; 1995-200120/26.

XX P-PSDB; AAR76238.

XX New nucleic acid encoding Fas protein without its trans-membrane region  
 and related vectors, transformed cells, transgenic animals, protein and  
 antibodies, useful for control of Fas mediated apoptosis

XX Claim 3; Fig.3-1 to 3-4; 38pp; English.

XX mRNA was obt'd. from human lymphocytes and PCR was used to make  
 cDNA specific for Fas-delta-TM (i.e. Fas lacking the transmembrane  
 region) mRNA. The PCR product was ligated into pBluescript and the  
 recombinant plasmid was used to transfect E. coli DH5-alpha cells. The  
 insert sequence of pBluescript-Fas-delta-TM is given in AAQ93879.  
 (updated on 25-MAR-2003 to correct PN field.)

XX Sequence 2471 BP; 807 A; 474 C; 489 G; 701 T; 0 other;

Query Match 7.0%; Score 226; DB 16; Length 2471;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-50;  
 Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2033 GACGCTTCGGGGAGTGAAGGAGCGTTTACAGTGAAGCTTGAGAGCTCAGGGGCG 2092  
 1 GACGCTTCGGGGAGTGAAGGAGCGTTTACAGTGAAGCTTGAGAGCTCAGGGGCG 60

DB 2093 GGCACCTGGACGAGAACACACCTCTGAGGCGCAGCCCTGGCTGCCAGGCGAGCTGCTCTT 2152  
 61 GGCACCTGGACGAGAACACACCTCTGAGGCGCAGCCCTGGCTGCCAGGCGAGCTGCTCTT 120

QY 2153 CTCCTGGCGGGTGTGTGAGACCCGCTCAGTACGAGAGTGGAGAGCTTTCACTTCGAGG 2212  
 121 CTCCTGGCGGGTGTGTGAGACCCGCTCAGTACGAGAGTGGAGAGCTTTCACTTCGAGG 180

DB 2213 ATTGCTCAACACCATGCTGGGCACTTGACCCCTCTCACTCTGGT 2258  
 181 ATTGCTCAACACCATGCTGGGCACTTGACCCCTCTCACTCTGGT 226

## RESULT 13

AAK24878 ID AAK24878 standard; DNA; 2471 BP.

XX AAK24878;

XX 21-JUN-1999 (first entry)

XX Soluble Fas receptor DNA.

XX Fas receptor; Fas ligand; FasL; proinflammatory; immunosuppressive;  
 graft versus host disease; autoimmune disease; psoriasis;  
 rheumatoid arthritis; systemic lupus erythematosus; gene therapy;  
 ss.

XX Mammalia.

XX Key Location/Qualifiers

FT CDS 195..1139

FT sig\_peptide /tag= a

FT mat\_peptide /tag= b

FT /tag= c

FT /tag= d

FT /tag= e

XX MO9903999-A1.

XX 28-JUN-1999.

XX 16-JUL-1998; 98WO-US14771.

XX 17-JUL-1997; 97US-0052829.

XX (UNMI ) UNIV MICHIGAN.

XX Chen J, Nabel GJ;

XX WPI; 1999-132243/11.

XX P-PSDB; AAW98070.

XX Inhibition of proinflammatory responses - using an agent which  
 modulates FasL stimulation, used for treating graft versus host  
 disease or autoimmune disease

XX Disclosure; Fig 4B; 71pp; English.

XX This present sequence is a DNA clone encoding soluble Fas receptor

XX

XX

XX

XX

XX

XX

XX

(see AAM98070). The invention provides a method for inhibiting a CC proinflammatory response in a cell mixture by administering an CC immunosuppressive agent which inhibits the proinflammatory activity of Fas ligand (FasL). In some embodiments, FasL is coadministered with the immunosuppressive agent, and the cell mixture comprises neutrophil cells. The method can be practiced in vitro, ex vivo or in vivo. Suitable immunosuppressive agents include antisense molecules that inhibit endogenous FasL expression, soluble Fas receptors, ribozymes that inhibit the endogenous expression of FasL, drugs that inhibit FasL signaling, agents that induce the CC and polynucleotides coding for an immunosuppressive agent such as TGF-beta. The method can be used for treating diseases associated with an undesired FasL-mediated proinflammatory response, e.g., CC graft versus host disease, or an autoimmune disease such as systemic lupus erythematosus, rheumatoid arthritis and psoriasis. The invention also provides a method for identifying agents which modulate FasL stimulation of a proinflammatory response.

Sequence 2471 BP; 807 A; 474 C; 489 G; 701 T; 0 other;

Query Match 7.0%; Score 226; DB 20; Length 2471;  
Best Local Similarity 100.0%; Pred. No. 1.5e-50;

Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2033 GACGCTTCGGGGAGTGAAGGAGCGTTTACGAGTGAAGTGGCTGAGAGCCCTCAGGGCG 2092  
DB 1 GACGCTTCGGGGAGTGAAGGAGCGTTTACGAGTGAAGTGGCTGAGAGCCCTCAGGGCG 60  
QY 2093 GGCACCTGCGACGGAACACACCTGAGGCGAGCCCTGCTGCCAGGGAGCTCCCTCTT 2152  
DB 61 GGCACCTGCGACGGAACACACCTGAGGCGAGCCCTGCTGCCAGGGAGCTCCCTCTT 120  
QY 2153 CTCCTGGGGGTTGGTGGAGCCCGCTCACTGAGTGGAGGAGCTTTCACTTCGGAGG 2212  
DB 121 CTCCTGGGGGTTGGTGGAGCCCGCTCACTGAGTGGAGGAGCTTTCACTTCGGAGG 180  
QY 2213 ATTGCTCAACCAACCATCTGGGAGCTGAGAGCCCTCACTTCGGT 2258  
DB 181 ATTGCTCAACCAACCATCTGGGAGCTGAGAGCCCTCACTTCGGT 226

RESULT 14  
AAQ29959 standard; cDNA to mRNA; 2534 BP.

AC AAQ29959;  
XX 25-MAR-2003 (updated)  
DT 12-MAR-1993 (first entry)  
XX Human cell surface antigen.  
DE Fas antigen; apoptosis; pF58; NGFR/TNFR family; ss.  
XX Homo sapiens.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT sig\_peptide /tag= a  
FT mat\_peptide /tag= b  
FT /product= Fas\_antigen  
FT conflict /tag= c  
FT /note= "this residue is not present in pf3"  
FT polyA\_signal /tag= d  
FT polyA\_signal /tag= e  
FT polyA\_signal /tag= e  
FT polyA\_signal /tag= e  
XX

PN EP510691-A1.  
XX 28-OCT-1992.  
XX  
PF 24-APR-1992; 92EP-0107060.  
XX  
PR 26-APR-1991; 91JP-0125234.  
XX  
PA (OSAB-) OSAKA BIOSCIENCE INST.  
PI Itoh N, Nagata S, Yonehara S;  
DR WPI, 1992-358914/44.  
DR P-PSDB; AAR28084.  
PT DNA encoding human cell surface antigen - used to clarify  
PT apoptosis mechanism of various types of cell, and to prepare  
XX monoclonal antibodies that react with tumour cells expressing Fas  
PS Claim 3; Fig 1 and 2; 27pp; English.

A cDNA library was prepared from polyA+ RNA from the human lymphoma cell line KT-3. The cDNA was ligated to BstXI-cut vector pCEV4 via BstXI linkers. The KT3 cDNA library was used to transfect monkey COS-7 cells which were then suspended in buffer containing murine anti-Fas antibodies. The Fas-expressing cells adhered to the plates. Extrachromosomal DNA was prepared from adhered cells and used to transform E.coli WMI00 cells. A 520bp XhoI-BamHI fragment from a positive clone (pf3) was used to screen the KT-3 cDNA library. The longest cDNA clone was designated pF58 and contains an ORF corresp. to a 335 amino acid pre-protein and a 319 amino acid mature protein (1.e. human Fas antigen).  
(Updated on 25-MAR-2003 to correct PN field.)

Sequence 2534 BP; 817 A; 487 C; 503 G; 727 T; 0 other;

Query Match 7.0%; Score 226; DB 13; Length 2534;  
Best Local Similarity 100.0%; Pred. No. 1.5e-50;

Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2033 GACGCTTCGGGGAGTGAAGGAGCGTTTACGAGTGAAGTGGCTGAGAGCCCTCAGGGCG 2092  
DB 1 GACGCTTCGGGGAGTGAAGGAGCGTTTACGAGTGAAGTGGCTGAGAGCCCTCAGGGCG 60  
QY 2093 GGCACCTGCGACGGAACACACCTGAGGCGAGCCCTGCTGCCAGGGAGCTCCCTCTT 2152  
DB 61 GGCACCTGCGACGGAACACACCTGAGGCGAGCCCTGCTGCCAGGGAGCTCCCTCTT 120  
QY 2153 CTCCTGGGGGTTGGTGGAGCCCGCTCACTGAGTGGAGGAGCTTTCACTTCGGAGG 2212  
DB 121 CTCCTGGGGGTTGGTGGAGCCCGCTCACTGAGTGGAGGAGCTTTCACTTCGGAGG 180  
QY 2213 ATTGCTCAACCAACCATCTGGGAGCTGAGAGCCCTCACTTCGGT 2258  
DB 181 ATTGCTCAACCAACCATCTGGGAGCTGAGAGCCCTCACTTCGGT 226

RESULT 15  
AAQ95297 standard; cDNA; 2534 BP.

AC AAQ95297;  
XX 19-FEB-1996 (first entry)  
DT 19-FEB-1996 (first entry)  
XX Plasmid pF58 contg. human Fas cDNA.  
DE Plasmid pF58; human Fas cDNA; soluble membrane protein;  
XX antibody production; diseases; treatment; prevention; ds.  
XX Homo sapiens.  
OS



```

FH Key Location/Qualifiers
FT CDS 195..1202
FT sig_peptide /+tag= a
FT sig_peptide 195..242
FT mat_peptide /+tag= b
FT mat_peptide 243..1199
FT /+tag= c
XX JP07115988-A.
XX
XX 09-MAY-1995.
XX
XX 26-OCT-1993; 93JP-0267644.
XX
XX 26-OCT-1993; 93JP-0267644.
XX
XX (NISB ) JAPAN TOBACCO INC.
XX
XX WPI; 1995-202847/27.
XX P-PSDB; AAR78606.
XX
XX Preparation of soluble membrane proteins - for their use in antibody
XX production for the treatment and prevention of related diseases
XX
XX Example 1; Pages 15-17; 51pp; Japanese.
XX
XX AAQ95297 is the plasmid pF58 which contains the human Fas CDNA. The
XX plasmid was used in the construction of an expression vector for
XX the prodn. of recombinant soluble membrane proteins. The proteins
XX can be used in antibody prodn. for the treatment and prevention of
XX related diseases.
XX
XX SQ Sequence 2534 BP; 817 A; 487 C; 503 G; 727 T; 0 other;
XX
Query Match 7.0%; Score 226; DB 16; Length 2534;
Best Local Similarity 100.0%; Pred. No. 1.5e-50;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2033 GACGCTTCTGGGAGTGAAGGAGCGGTTTACAGTGAAGTCTTGGCTGGAGCCCTCAGGGCG 2092
DB 1 GACGCTTCTGGGAGTGAAGGAGCGGTTTACAGTGAAGTCTTGGCTGGAGCCCTCAGGGCG 60
QY 2093 GGCACCTGGACGGAACACACCTCTGAGCCAGCCCTGGCTGCGCCAGCGAGCTGCTCTT 2152
DB 61 GGCACCTGGACGGAACACACCTCTGAGCCAGCCCTGGCTGCGCCAGCGAGCTGCTCTT 120
QY 2153 CTCCCGGCGGTTGTGAGACCCGCTCACTAGTACGAGTTGGGAGCTCTTCACTTGGAGG 2212
DB 121 CTCCCGGCGGTTGTGAGACCCGCTCACTAGTACGAGTTGGGAGCTCTTCACTTGGAGG 180
QY 2213 ATTGCTACACACCATGCTGAGCATCTTGGACCTCTTCACTTGGT 2258
DB 181 ATTGCTACACACCATGCTGAGCATCTTGGACCTCTTCACTTGGT 226

```

Search completed: August 1, 2003, 13:55:11  
Job time : 1079.81 secs



Run on: August 1, 2003, 08:03:17 (without alignments)  
7955.924 Million cell updates/sec

```

Title:      US-09-834-291.1
Perfect score: 3212
Sequence:   1  ttgaggactctcagagatatg.....ttgcagatgctcaatcaag 3212

```

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1455/07 2097 2879534  
 - Choosing chosen parameters:

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
                  Maximum Match 100%

```

Maximum Listing	first 45 summaries
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

```
Database : Pubmed
1: /cgn2_6/ptodata/1/pubpna/PCT_NEW PUB. seq:*
```

```

2:      /cgn2_6/ptodata/1/pubnba/US06_NEW_PUB.seq.*
3:      /cgn2_6/ptodata/1/pubnba/US06_NEW_PUB.seq.*
4:      /cgn2_6/ptodata/1/pubnba/US07_NEW_PUB.seq.*
5:      /cgn2_6/ptodata/1/pubnba/PC07_PUBCOMB.seq.*
6:      /cgn2_6/ptodata/1/pubnba/US08_NEW_PUB.seq.*
7:      /cgn2_6/ptodata/1/pubnba/US08_PUBCOMB.seq.*
8:      /cgn2_6/ptodata/1/pubnba/US08A_PUBCOMB.seq.*
9:      /cgn2_6/ptodata/1/pubnba/US09C_PUBCOMB.seq.*
10:     /cgn2_6/ptodata/1/pubnba/US09C_PUBCOMB.seq.*
11:     /cgn2_6/ptodata/1/pubnba/US09C_PUBCOMB.seq.*
12:     /cgn2_6/ptodata/1/pubnba/US10A_PUBCOMB.seq.*
13:     /cgn2_6/ptodata/1/pubnba/US10B_PUBCOMB.seq.*
14:     /cgn2_6/ptodata/1/pubnba/US10_NEW_PUB.seq.*
15:     /cgn2_6/ptodata/1/pubnba/US60_NEW_PUB.seq.*
16:     /cgn2_6/ptodata/1/pubnba/US60_PUBCOMB.seq.*
17:

```

1.  $\chi^2$  -  $\chi^2$  test. Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3212	100.0	3212	9	US-09-834-291-1	Sequence 1, Appl1
2	2346	73.1	2837	9	US-09-834-291-4	Sequence 94, Appl1
3	2143	66.7	2165	9	US-09-802-669-94	Sequence 3, Appl1
4	1899.6	59.1	2380	9	US-09-834-291-3	Sequence 2, Appl1
5	713	22.2	720	9	US-09-834-291-2	Sequence 167, App
6	2874.4	8.9	859	9	US-09-923-299-187	Sequence 187, App
7	287.4	8.9	859	11	US-09-923-299-187	Sequence 32, Appl1
8	262.8	8.2	266	9	US-09-834-291-32	Sequence 1, Appl1
9	252	7.8	2551	9	US-09-802-669-1	Sequence 16, Appl1
10	226	7.0	2534	10	US-09-949-713-16	Sequence 1, Appl1
11	60.8	1.9	816	13	US-09-884-987-1	Sequence 145670
12	60	1.9	816	10	US-09-926-632-145670	Sequence 796, App
13	60	1.9	2435	10	US-09-902-941-796	Sequence 796, App
14	60	1.9	2435	10	US-09-902-941-796	Sequence 796, App
15	60	1.9	2435	10	US-09-849-626-796	Sequence 796, App
16	60	1.9	2435	14	US-10-017-754-796	Sequence 796, App

17	60	1.9	2821	10	US-09-736-457-1669	Sequence 1669, AP
C 18	60	1.9	2821	10	US-09-902-941-1669	Sequence 1669, AP
C 19	60	1.9	2821	10	US-09-849-626-1669	Sequence 1669, AP
C 20	60	1.9	2821	14	US-10-017-754-1669	Sequence 91808, A
21	59.2	1.8	542	13	US-10-027-632-91808	Sequence 304597, A
22	59.2	1.8	542	13	US-10-027-632-904597	Sequence 134015, A
C 23	59.2	1.8	639	13	US-10-027-632-134014	Sequence 134015, A
C 24	59.2	1.8	639	13	US-10-027-632-134015	Sequence 134016, A
C 25	59.2	1.8	639	9	US-10-027-632-134016	Sequence 41, Appl
C 26	59.2	1.8	8888	13	US-09-836-077-41	Sequence 3, Appl
C 27	58.4	1.8	9947	14	US-10-251-857-3	Sequence 989, Appl
28	57	1.8	975	14	US-09-954-531-998	Sequence 103, App
C 29	56.2	1.7	2892	14	US-10-175-532-103	Sequence 253750, A
30	55.8	1.7	1140	13	US-10-027-632-253750	Sequence 253751, A
31	55.2	1.7	443	13	US-10-027-632-253751	Sequence 64316, A
32	55.2	1.7	443	13	US-10-027-632-64316	Sequence 64317, A
33	55.2	1.7	443	13	US-10-027-632-296990	Sequence 296990, A
34	55.2	1.7	443	13	US-10-027-632-296991	Sequence 296991, A
35	55.2	1.7	443	13	US-10-067-514-1	Sequence 1, Appl
36	54.8	1.7	1691139	14	US-10-027-632-103140	Sequence 103140, A
37	54.6	1.7	2501	13	US-10-027-632-103140	Sequence 3, Appl
C 38	54.6	1.7	4876	14	US-10-284-048-3	Sequence 3, Appl
39	54.6	1.7	174493	10	US-09-804-471A-3	Sequence 71673, A
40	54.6	1.7	174493	14	US-10-238-709-3	Sequence 71673, A
41	54.4	1.7	551	13	US-10-027-632-71673	Sequence 11168, A
42	54	1.7	2759	13	US-10-027-632-11168	Sequence 84928, A
43	53.8	1.7	478	13	US-10-027-632-84929	Sequence 84930, A
44	53.8	1.7	478	13	US-10-027-632-84930	Sequence 84930, A
45	53.4	1.7	572	13	US-10-027-632-238591	Sequence 238591, A

## ALIGNMENTS

RESULT 1  
MS-09-834-291-1  
MS/09834291

, Sequence 1, applicant: 109768  
Patent No. US20020042064N1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPlicant: Muller-Schilling, Martina  
Applicant: Oren, Moshe  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
PRIOR FILING DATE: 2001-08-21  
PCT/PCT/DE99/03343  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1999-10-18  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 3212  
TYPE: DNA  
ORGANISM: Homo Sapiens

US-09-834-291-1

w

[illegible]





841 GCGTCGACAGCCAGGGCTTCTCTCATGCGCACTAAGCTACTGAAAGGTGAAACAGAG 900  
QY 901 ACAAGCCCTATCAACCTTACAGAGCTGTGTGTAGTACGATGACAGATGCAAAACAGAG 960  
Db 901 ACAAGCCCTATCAACCTTACAGAGCTGTGTGTAGTACGATGACAGATGCAAAACAGAG 960  
QY 961 GTATGAAAGCCCTCAGAGAGGTAACTTACCTTGAATTTAGAGGCC-ACAGGCTCCA 960  
Db 961 GTATGAAAGCCCTCAGAGAGGTAACTTACCTTGAATTTAGAGGCC-ACAGGCTCCA 960  
QY 1020 GAAGAAATATGCACTGAGAGAGAGCTGAAAGATGAAAGTGGCTTACGAAAGGTAA 1019  
Db 1021 GAAGAAATATGCACTGAGAGAGAGCTGAAAGATGAAAGTGGCTTACGAAAGGTAA 1079  
QY 1081 TTAATGTGTTTAAATGAGGTGATCTAATTTGGGAGAGAGAGAGGTGCAAGAGGTAA 1080  
Db 1081 TTAATGTGTTTAAATGAGGTGATCTAATTTGGGAGAGAGAGAGGTGCAAGAGGTAA 1080  
QY 1140 GCAGAGCTGTGTGAGAGAGAGCTGAAAGATGAAAGTGGCTTACGAAAGGTAA 1140  
Db 1141 GCAGAGCTGTGTGAGAGAGAGCTGAAAGATGAAAGTGGCTTACGAAAGGTAA 1199  
QY 1200 CTGCAATCAAAATTCAGAGTCAATATGATGATCAATATCAAAAGTCTGTAAAT 1200  
Db 1201 CTGCAATCAAAATTCAGAGTCAATATGATGATCAATATCAAAAGTCTGTAAAT 1259  
QY 1260 CATGCTAAATTAACCTTAAGAGCTATCTAAGCTTCAAGAGAGAGAGAGAGAGGT 1260  
Db 1261 CATGCTAAATTAACCTTAAGAGAGCTATCTAAGCTTCAAGAGAGAGAGAGAGGT 1319  
QY 1320 TCACCAAGGT 1320  
Db 1321 TCACCAAGGT 1379  
QY 1380 GAGTAAAGGT 1380  
Db 1381 GAGTAAAGGT 1439  
QY 1440 GCTTAAAGTGTAGAGT 1440  
Db 1441 GCTTAAAGT 1499  
QY 1500 CAGAGCCCTTATGAG 1500  
Db 1501 CAGAGCCCTTATGAG 1559  
QY 1560 GTCTGTAGAGCTCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1560  
Db 1561 GTCTGTAGAGCTCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1619  
QY 1620 TCTTTCTGTAG 1620  
Db 1621 TCTTTCTGTAG 1679  
QY 1680 GCTGTAGAGCTCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1680  
Db 1681 GCTGTAGAGCTCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1739  
QY 1740 GCTTTTGAAG 1740  
Db 1741 GCTTTTGAAG 1799  
QY 1800 CACTGTAG 1800  
Db 1801 CACTGTAG 1859  
QY 1860 TCAACCTGTCTTCT 1860  
Db 1861 TCAACCTGTCTTCT 1919  
QY 1920 CTTCCCAAG 1920  
Db 1921 CTTCCCAAG 1979  
QY 1980 TTTTGTGAGAGCTGTCTCTATCTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2039  
Db 1981 TTTTGTGAGAGCTGTCTCTATCTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2039  
QY 2040 CTGGAG 2040  
Db 2041 CTGGAG 2099  
QY 2100 GCAG 2100  
Db 2101 GCAG 2159  
QY 2160 GGTGTGTGAG 2160  
Db 2161 GGTGTGTGAG 2219  
QY 2220 AACAG 2219  
Db 2221 AACAG 2279  
QY 2280 GAGCTTACCTGCTGT 2187  
Db 2281 GAGCTTACCTGCTGT 2187  
QY 2340 GGAATGTGAG 2209  
Db 2341 GGAATGTGAG 2209  
QY 2400 GTTGAAG 2209  
Db 2401 GTTGAAG 2259  
QY 2460 ATGAG 2209  
Db 2461 ATGAG 2259  
QY 2520 GCAAGCTGCGAG 2209  
Db 2521 GCAAGCTGCGAG 2259  
QY 2580 ACAG 2209  
Db 2581 ACAG 2259  
QY 2640 AG 2247  
Db 2641 AG 2299  
QY 2700 CCGGTGTGAG 2307  
Db 2701 CCGGTGTGAG 2359  
QY 2760 ACATATGTCTATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2367  
Db 2761 ACATATGTCTATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2419  
QY 2820 GGTAAAG 2427  
Db 2821 GGTAAAG 2479  
QY 2880 GGTAAAG 2487  
Db 2881 GGTAAAG 2539  
QY 2940 GCTGTGTCT 2547  
Db 2941 GCTGTGTCT 2599  
QY 3000 CATTTTGT 2607  
Db 3001 CATTTTGT 2659  
QY 3060 CATTTTGT 2667  
Db 3061 CATTTTGT 2667

QY	867	TGGCACTAACAGTCTACGAAAGGTGGAAAGAGACAGACAGCTTACACCTACCAAGCT	926
Db	421	TGGCACTAACAGTCTACGAAAGGTGGAAAGAGACAGACAGCTTACACCTACCAAGCT	480
QY	927	GGTGTAAAGTGCAGTGCAGATGCMAAACACAGGGTGTAGAAAAGCCCTCAGAGGGTAA	986
Db	481	GGTGTAAAGTGCAGTGCAGATGCMAAACACAGGGTGTAGAAAAGCCCTCAGAGGGTAA	540
QY	987	CCTAACCTAATTTGAGGGGCC-ACAGGCTCCAGAGAAATGTCAATCGAGAGAGC	1045
Db	541	CCTAACCTAATTTGAGGGGCCCAACAGGCTCCAGAGAAATGTCAATCGAGAGAGC	600
QY	1046	CTGAAGATGAACAGTGGGCTAGACAAAGGTTATATGTGTATTAATGGGTGAATC	1105
Db	601	CTGAAGATGAACAGTGGGCTAGACAAAGGTTATTAATGTGTATTAATGGGTGAATC	660
QY	1106	TAAATGGAGAGAGAGAGGTTGCAGATGAGGTTGCAGAGCTTGTGTAGCAATGCCAAG	1165
Db	661	TAAATGGAGAGAGAGAGGTTGCAGATGAGGTTGCAGAGCTTGTGTAGCAATGCCAAG	720
QY	1166	GAATACTGAACCTTAGTGTGTCCAGTCTGGAACCTGCAATCCAAATTCAGTTCAGTAAT	1225
Db	721	GAATACTGAACCTTAGTGTGTCCAGTCTGGAACCTGCAATCCAAATTCAGTTCAGTAAT	780
QY	1226	GATGTCAATTATCCAAACATACCTTCTGTAAATTCATGTAACTACTCTTAAGGCTATCT	1285
Db	781	GATGTCAATTATCCAAACATACCTTCTGTAAATTCATGTAACTACTCTTAAGGCTATCT	840
QY	1286	ACCGTCCAAAGCAATGTGACTTTGAAACAGTGTCAACAGAGCAGMAAGATTACAG	1345
Db	841	ACCGTCCAAAGCAATGTGACTTTGAAACAGTGTCAACAGAGCAGMAAGATTACAG	900
QY	1346	ATTTTATTTTAAAGAAATTTGACAGAAATTAATGATACAGAGACAGAAAGTAATG	1405
Db	901	ATTTTATTTTAAAGAAATTTGACAGAAATTAATGATACAGAGACAGAAAGTAATG	960
QY	1406	TGAATGTTAATATATGCTGGGGCTATCGAATTTGGCTTAAGTTGTTGTTTCT	1465
Db	961	TGAATGTTAATATATGCTGGGGCTATCGAATTTGGCTTAAGTTGTTGTTTCT	1020
QY	1466	CTTGAGAAATTAATACTAAAGGGCTCTCTTTCAAGCTTATGAGCGCAATCTGTA	1525
Db	1021	CTTGAGAAATTAATACTAAAGGGCTCTCTTTCAAGCTTATGAGCGCAATCTGTA	1080
QY	1526	CTTTTTCATATGTTAATGCTCATTTCCAGAAAGCTGTGAGGCTCTCATGTTGCAAGC	1585
Db	1081	CTTTTTCATATGTTAATGCTCATTTCCAGAAAGCTGTGAGGCTCTCATGTTGCAAGC	1140
QY	1586	ACAACATGACAGCCCAAGTCAATGACCCGCAAGCTTCTCTGATGATCTCCAGCAAT	1645
Db	1141	ACAACATGACAGCCCAAGTCAATGACCCGCAAGCTTCTCTGATGATCTCCAGCAAT	1200
QY	1646	AGCCAAAGGCTCTGTAACCCAGAGAGCCTGTGGCTGTGAAGTCCATCTCTTCAAGA	1705
Db	1201	AGCCAAAGGCTCTGTAACCCAGAGAGCCTGTGGCTGTGAAGTCCATCTCTTCAAGA	1260
QY	1706	CCCTCCCACTTCCAGAGTGAACATACAGCAGAAAGCTTTAAGAAAGGACAGAGGCGGC	1765
Db	1261	CCCTCCCACTTCCAGAGTGAACATACAGCAGAAAGCTTTAAGAAAGGACAGAGGCGGC	1320
QY	1766	TCTGAGAGTCTCACTGAATGACATGACAGCAGCAGCAGAGAGCGCCCGGACAGGA	1825
Db	1321	TCTGAGAGTCTCACTGAATGACATGACAGCAGCAGCAGAGAGCGCCCGGACAGGA	1380
QY	1826	ATGGCCATTTGTGCAAGAAACCTGACTCTTCTTCAACCTGACTCTTCTCCCTCTAC	1885
Db	1381	ATGGCCATTTGTGCAAGAAACCTGACTCTTCTTCAACCTGACTCTTCTCCCTCTAC	1440
QY	1886	CGGCGGACGCGCAAGTTGCTGAATCAATGAGCCCTCCCAACCGGCGGTCTCCAGC	1945
Db	1441	CGGCGGACGCGCAAGTTGCTGAATCAATGAGCCCTCCCAACCGGCGGTCTCCAGC	1500
QY	1946	GAGGCTTCTTCCATCTCTCTGACACCGGAGGCTTTGTGTGAGCTGTCTGATCTCG	2005

conservative; Mismatch

1 AGCTTTTGGCTACA TTTTTTTT TTTTTCATGAGTAAATCATCATTCACTG 507

61 GGGTATAATCTGTTCTGAGTTGCTGGCTTATAT 567

[illegible]

|||||CCCTTACCTCTCCTTCCTTCCCTCACACCCCCTTTTCCCTGCGGG

086 CTTTACATTTTTATTAAATGAAGCTTTTCATTTTGGAATACTGTAAGCCTTT

748 AAATTTCAGAGATAATACAGAGAATGCCCATATACCTCCGCGTGGTT  
-----CAGCAAAACA 300

808 TGTGCTATTAGATGCTCAGAGTGTCTGCA... 360

868 GGCACCTACAGTCTACTGATTCCTCAT 420

928 GTGTTAAGTCCACCTCCTATCAACAACCTACAAGACTG 480

988 CMT3A cDNA 540

.....CCAGAAITGAGGGCCCAACAGGCTCCAGAGAAATGTCAACTGAGAGGCAACCG

801 1GAGGATGAACACTGGCTAAGCAAGGTTATTAAATGTGTTATTTTTCCTTTT  
1200

661 AATTGGAGCGAGAGGTTGCAGACTGAGCTCAGCCACCAAGG 1166

721 AATACTGAACCTTTAGTGTGTCCTACTCCTCAATGGCGAATGCTGCAGGTCAGTAATG 1226

781 ATGTCATTATCCAAACATAGCTTCTA 1286

841 CCGTTCCGAGACCACTTGGATGTCACCAAGCAGAAAGAATTACAAGA 1346

907  
|||||  
|||GCGAGGAATATGAGTAACGAGGACACAGGAGTAAATTGT 1405

|||||ATTAHAGCTGGGCTATGCCATTGGCTTAA GTTGTTAGCTTTGTTCCTCCCGC

11GAGGAATAAAACTAAGGGCCCTCCCTTTTCAGAGCCTTATGCCCCA...

1



```

Db      1021 TTGAGAAATAAAACTAAGGGGCGCTTCCTTTTCAGAGCCCTATGGCGCAACATCTGTAC 1080
Qy      1527 TTTTTCATANGTTAACTGTCCATTTCCAGAAAGCTCTGTGAAGCTCTCATGTGTGACGCA 1586
Db      1081 TTTTTCATANGTTAACTGTCCATTTCCAGAAAGCTCTGTGAAGCTCTCATGTGTGACGCA 1140
Qy      1587 CAACATGACAGCCGACATCAATGATCCCGGCAAGCTTTCTCTGAGTACTCCAGCAATTA 1646
Db      1141 CAACATGACAGCCGACATCAATGATCCCGGCAAGCTTTCTCTGAGTACTCCAGCAATTA 1200
Qy      1647 GCCAAGGCTCTGTATCCAGAGGAGCACTCTGCGCTCTGAGCTCATTTCTCTTCAAGAC 1706
Db      1201 GCCAAGGCTCTGTATCCAGAGGAGCACTCTGCGCTCTGAGCTCATTTCTCTTCAAGAC 1260
Qy      1707 CTCCCAACTTCCAGGTTGAACTACAGCAAGGCTTTAGAAAGGAGGAGGAGGCGGCT 1766
Db      1261 CTCCCAACTTCCAGGTTGAACTACAGCAAGGCTTTAGAAAGGAGGAGGAGGCGGCT 1320
Qy      1767 CTGAGGCTCTCACTGAAAGTGAAGTCCAGCACTGAGAGAGGCGGCGGAGAGGAA 1826
Db      1321 CTGAGGCTCTCACTGAAAGTGAAGTCCAGCACTGAGAGAGGCGGCGGAGAGGAA 1380
Qy      1827 TGCCCATTTGTGACAGAAAGCTGTGACTCTTCTCTCACTGACTTCTCTCTCTCTTACC 1886
Db      1381 TGCCCATTTGTGACAGAAAGCTGTGACTCTTCTCTCACTGACTTCTCTCTCTCTTACC 1440
Qy      1887 GCGCGCAGGCGCAAGTTGCTGAATCATGAGAGGCTCTCCCAAGGCGGCGGCTCTCCAGCG 1946
Db      1441 GCGCGCAGGCGCAAGTTGCTGAATCATGAGAGGCTCTCCCAAGGCGGCGGCTCTCCAGCG 1500
Qy      1947 AGGCTTCTTCCCATCTCTCTGACCAAGCGGCGCTTTTGTGAGCTCGTCTCTGATCTCGC 2006
Db      1501 AGGCTTCTTCCCATCTCTCTGACCAAGCGGCGCTTTTGTGAGCTCGTCTCTGATCTCGC 1560
Qy      2007 GCAAGAGTACACACAGTGTTCAAAGAGCTTCTGAGGAGTGAAGGAGGAGGAGGAGGAGG 2066
Db      1561 GCAAGAGTACACACAGTGTTCAAAGAGCTTCTGAGGAGTGAAGGAGGAGGAGGAGGAGG 1620
Qy      2067 GTGACTTGTGAGAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2126
Db      1621 GTGACTTGTGAGAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
Qy      2127 TGAGTGGCCAGGCGGAGCTGCTCTTCTCCCGGGGAGTGGTGGAGCCCGCTCACTAGAGAG 2186
Db      1681 TGAGTGGCCAGGCGGAGCTGCTCTTCTCCCGGCGGAGTGGTGGAGCCCGCTCACTAGAGAG 1727
Qy      2187 TTTGGGAGGCTTTTCACTTCTGAGAGATGCTCAACAACATGCTGGGAGCTGAGACCT 2246
Db      1728 CTGAGAGATCT----- 1740
Qy      2247 CCTACCTGTGATCCTCTCTGCGCGGAGGAGGCTTACCCCGTCTTAGTCCGCGG 2306
Db      1741 ----- 1740
Qy      2307 ATTAGCAAGTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2366
Db      1741 AGTGGCAAGTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1765
Qy      2367 GGCACCTGGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2426
Db      1766 ----- 1765
Qy      2427 AGGACCTTCCCTAGGCGCGGAGTGTCTAGAAAGATGAGAGACTGTCTTTCTTGGGCGCT 2486
Db      1766 ----- 1765
Qy      2487 GATGCAAGTGTCTATCCCGTGGGAGAGGCGGAGCACTCCGCGCTCTCGAGAGCAC 2546
Db      1766 ----- 1765
Qy      2547 TGCGCTCAAGTTGAGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2606
Db      1766 -----AA 1767

```

```

Qy      2607 GCTTAAAGTGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2666
Db      1768 GCTTAAAGTGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1827
Qy      2667 CTTGACAAAGCAAGCCAAAGGTCGCTCCGCGCGGAGTGGAGTGGAGGAGGAGGAGGAGGAGG 2726
Db      1828 CTTGACAAAGCAAGCCAAAGGTCGCTCCGCGCGGAGTGGAGTGGAGGAGGAGGAGGAGGAGG 1887
Qy      2727 GGGGCGGAGAGAGGCTCAAGGCTTCAAGACATATTTGCTCATTTTCTGGCAATTTCT 2786
Db      1888 GGGGCGGAGAGAGGCTCAAGGCTTCAAGACATATTTGCTCATTTTCTGGCAATTTCT 1947
Qy      2787 CAGAGTGAAGAAATAGTCAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2846
Db      1948 CAGAGTGAAGAAATAGTCAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2007
Qy      2847 GCACTTTTCTTCTGAAAAAGTATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2906
Db      2008 GCACTTTTCTTCTGAAAAAGTATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2067
Qy      2907 ACCGTTTTTATTTGACACAGAAAGAAAGTGTCTCTCTCCGCGGAGGAGGAGGAGGAGG 2966
Db      2068 ACCGTTTTTATTTGACACAGAAAGAAAGTGTCTCTCTCTCCGCGGAGGAGGAGGAGGAGG 2127
Qy      2967 CTTTAAAGCTGTAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3026
Db      2128 CTTTAAAGCTGTAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2187
Qy      3027 CTTTCTTCTTTTGGCTTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3086
Db      2188 CTTTCTTCTTTTGGCTTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2247
Qy      3087 CTGGGAGTGTGATAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3146
Db      2248 CTGGGAGTGTGATAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2307
Qy      3147 ACTTGAAGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3206
Db      2308 ACTTGAAGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2367
Qy      3207 TCAAG 3212
Db      2368 TCAAG 2373

```

RESULT 5  
US-09-834-291-2  
Sequence 2, Application US/09834291  
Patent No. US20020042064A1

GENERAL INFORMATION: Peter  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1998-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 720  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-2

Query Match 22.2%; Score 713; DB 9; Length 720;  
Best Local Similarity 100.0%; Pred. No. 3.3e-186;  
Matches 713; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 2500 GATCCCGCTGGGCGAGCGGCGAGCTCCGCGCTCTCGAGACCACTGGGCTTCACGTT 2559
DB 1 GATCCCGCTGGGCGAGCGGCGAGCTCCGCGCTCTCGAGACCACTGGGCTTCACGTT 60
QY 2560 GAGTGGGCGCTGGGCGAGCGGCGAGCTTCGAGAACTGCTGGGAACTTTAGGGTCCG 2619
DB 61 GAGTGGGCGCTGGGCGAGCGGCGAGCTTCGAGAACTGCTGGGAACTTTAGGGTCCG 120
QY 2620 TGGAGGCGGAGCGGCGAGCTTCGAGAACTGCTGGGAACTTTAGGGTCCG 2679
DB 121 TGGAGGCGGAGCGGCGAGCTTCGAGAACTGCTGGGAACTTTAGGGTCCG 180
QY 2680 GCCAAAGGCTCCGCTCCGCGCGGCTGGGCTGGGCGGCGGCGGCGGCGGCGGCGG 2739
DB 181 GCCAAAGGCTCCGCTCCGCGCGGCTGGGCTGGGCGGCGGCGGCGGCGGCGGCGG 240
QY 2740 GAGCTTACAGCTTCAGAACATATTTGCTCATTTTCTGGGAGTTCTGAGAGTAA 2799
DB 241 GAGCTTACAGCTTCAGAACATATTTGCTCATTTTCTGGGAGTAA 300
QY 2800 TAAGTCAGCACCGGAGAGCTGTTAAGCCGAGGCGTCCGAGAGACCGCACCTTTTCTT 2859
DB 301 TAAGTCAGCACCGGAGAGCTGTTAAGCCGAGGCGTCCGAGAGACCGCACCTTTTCTT 360
QY 2860 CTCGAAAAGTTATATGAGGCGGCTGAATGAGCTTCTGGAGCTTGTATACGTTTAT 2919
DB 361 CTCGAAAAGTTATATGAGGCGGCTGAATGAGCTTCTGGAGCTTGTATACGTTTAT 420
QY 2920 GTCAACAGAAAAGAACTGCTTCTCCCTCCGAGAACTTCTCTTAAAGCTGTA 2979
DB 421 GTCAACAGAAAAGAACTGCTTCTCCCTCCGAGAACTTCTCTTAAAGCTGTA 480
QY 2980 AGTCGCTGCTGAGTGTGTTATTTGTTTGTGTTTCTGCTTCTCTTCTTCTT 3039
DB 481 AGTCGCTGCTGAGTGTGTTATTTGTTTGTGTTTCTGCTTCTCTTCTTCTT 540
QY 3040 CCCTTTCTAGCTTCACATCCCATGCTGATTTCTGCTGCTGCTGCTGCTGCTGCTG 3099
DB 541 CCCTTTCTAGCTTCACATCCCATGCTGATTTCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 3100 TACTGCTTCCACCGCAGACAGAACCGGCGCTATTTATTTGGCAAACTTGAAGACCT 3159
DB 601 TACTGCTTCCACCGCAGACAGAACCGGCGCTATTTATTTGGCAAACTTGAAGACCT 660
QY 3160 GTTTGAAAAGTCCCTGCTCAGAAATGCCAGCTTGCAGATGCTTAATCAAG 3212
DB 661 GTTTGAAAAGTCCCTGCTCAGAAATGCCAGCTTGCAGATGCTTAATCAAG 713

```

## RESULT 6

```

US-09-925-299-187
; Sequence 187, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; FILE REFERENCE: Nucleic Acids, Proteins and Antibodies
; CURRENT APPLICATION NUMBER: US/09/925,299
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 187
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (803)

```

```

; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (853)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-187

```

```

Query Match
Best Local Similarity 8.9%; Score 287.4; DB 9; Length 859;
Matches 294; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

```

QY 1954 CTTCACATCTCTGACACACCGGAGCTTTCTGAGCTGCTCTGATCTCGGCAAGAG 2013
DB 19 CTTCACATCTCTGACACACCGGAGCTTTCTGAGCTGCTCTGATCTCGGCAAGAG 78
QY 2014 TGACACACAGGTGTTAAAGAGCGCTTCTGGGAGTGAAGGAGCGGTTTACAGAGTACT 2073
DB 79 TGACACACAGGTGTTAAAGAGCGCTTCTGGGAGTGAAGGAGCGGTTTACAGAGTACT 138
QY 2074 GGTGAGGCTTCAGGCGGCGGCACTGGCAACGACACCTTGAAGCCAGCTTGGCTGC 2133
DB 139 GGTGAGGCTTCAGGCGGCGGCACTGGCAACGACACCTTGAAGCCAGCTTGGCTGC 198
QY 2134 CCAGCGGAGCTGCTCTTCTCCGCGGTTGTGAGACCGCTCAGTACGAGTTGGGGA 2193
DB 199 CCAGCGGAGCTGCTCTTCTCCGCGGTTGTGAGACCGCTCAGTACGAGTTGGGGA 258
QY 2194 AGCTTTTCACTTGGAGATGCTCAACACATGCTGGCATCTGAGCTTCTTACT 2253
DB 259 AGCTTTTCACTTGGAGATGCTCAACACATGCTGGCATCTGAGCTTCTTACT 318
QY 2254 CTGCT 2258
DB 319 CTGCT 323

```

## RESULT 7

```

US-09-925-299-187
; Sequence 187, Application US/09925299
; Patent No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; FILE REFERENCE: Nucleic Acids, Proteins and Antibodies
; CURRENT APPLICATION NUMBER: US/09/925,299
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 187
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (803)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (853)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-187

```

```

Query Match
Best Local Similarity 8.9%; Score 287.4; DB 11; Length 859;
Matches 294; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

```

QY 1954 CTTCACATCTCTGACACACCGGAGCTTTCTGAGCTGCTCTGATCTCGGCAAGAG 2013
DB 19 CTTCACATCTCTGACACACCGGAGCTTTCTGAGCTGCTCTGATCTCGGCAAGAG 78
QY 2014 TGACACACAGGTGTTAAAGAGCGCTTCTGGGAGTGAAGGAGCGGTTTACAGAGTACT 2073

```

Db 79 TGACACACAGGCTTCAAAAGAGCTTCTGGGGAGTGAAGGAGCGGTTTACAGTACTT 138  
Qy 2074 GCGTGAAGCTTCAAGGGGCGGGGACCTGAGACACACCTTGAAGCCAGCCCTGGCTGC 2133  
Db 139 GCGTGAAGCTTCAAGGGGCGGGGACCTGAGACACACCTTGAAGCCAGCCCTGGCTGC 198  
Qy 2134 CCAGGCGGAGCTGCTCTTCTCCCGGGGTTGTGAGACCCGCTCACTGATGAGATTGGGGA 2193  
Db 199 CCAGGCGGAGCTGCTCTTCTCCCGGGGTTGTGAGACCCGCTCACTGATGAGATTGGGGA 258  
Qy 2194 AGCTCTTCACTTGGAGAGATTGCTCAACAACAGCTGAGATCTGAGACCTCTTACT 2253  
Db 259 AGCTCTTCACTTGGAGAGATTGCTCAACAACAGCTGAGATCTGAGACCTCTTACT 318  
Qy 2254 CTGCT 2258  
Db 319 CTGCT 323

RESULT 8  
US-09-834-291-32  
Sequence 32, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Kramer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT FILING DATE: US/09/834,291  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 32  
LENGTH: 266  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-32

Query Match 8.2%; Score 262.8; DB 9; Length 266;  
Best Local Similarity 99.2%; Pred. No. 4,6e-62;  
Matches 264; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 2500 GATCCGCTGGGCGAGGGGCGAGCTCCGGGCTCTCGAGACCACTGGCTCCAGCTT 2559  
Db 1 GATCCGCTGGGCGAGGGGCGAGCTCCGGGCTCTCGAGACCACTGGCTCCAGCTT 60  
Qy 2560 GAGGTGGGCTGGGCGAGGGGCGAGAAATTGAAGCGGAAGTCTGGGAAGCTTTAGGGTGGC 2619  
Db 61 GAGGTGGGCTGGGCGAGGGGCGAGAAATTGAAGCGGAAGTCTGGGAAGCTTTAGGGTGGC 120  
Qy 2620 TGAAGGGGAGCCCGGTTGAGAGAGAGAGCGGAATCTGTGACAAGCCCTTGAAGCCAA 2679  
Db 121 TGAAGGGGAGCCCGGTTGAGAGAGAGAGCGGAATCTGTGACAAGCCCTTGAAGCCAA 180  
Qy 2680 GCGAAGGCTGCTCCGGGCGAGGAGTGGTGAAGTGGGCGCCCGCCCGCGGGGCGGGGAGA 2739  
Db 181 GCGAAGGCTGCTCCGGGCGAGGAGTGGTGAAGTGGGCGCCCGCCCGCGGGGCGGGGAGA 240  
Qy 2740 GAGCTTACAGCTTTCAGAACATAT 2765  
Db 241 GAGCTTACAGCTTTCAGAACATAT 266

RESULT 9  
US-09-802-669-1  
Sequence 1, Application US/09802669  
Patent No. US20020004490A1

GENERAL INFORMATION:  
APPLICANT: Dean, Nicholas M.  
APPLICANT: Marcussen, Eric G.  
APPLICANT: Wyatt, Jacqueline  
APPLICANT: Zhang, Hong  
TITLE OF INVENTION: Antisense Compound Modulation of Fas Mediated Signaling  
FILE REFERENCE: 18PH-545  
CURRENT FILING DATE: US/09/802,669  
PRIOR FILING DATE: 2001-03-09  
PRIOR APPLICATION NUMBER: US 09/665,615  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: US 09/290,640  
PRIOR FILING DATE: 1999-04-12  
NUMBER OF SEQ ID NOS: 180  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 1  
LENGTH: 2551  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (221)..(1228)  
PUBLICATION INFORMATION:  
JOURNAL: J. Biol. Chem.  
VOLUME: 267  
ISSUE: 15  
PAGES: 10709-10715  
DATE: 1992-05-25  
DATABASE ACCESSION NUMBER: X63717/Genbank  
DATABASE ENTRY DATE: 1996-07-19  
US-09-802-669-1

Query Match 7.8%; Score 252; DB 9; Length 2551;  
Best Local Similarity 100.0%; Pred. No. 2.4e-58;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2007 GCAAGTGAACACAGAGTGTTCAAAGCGCTTCTGGGAGTGAAGGAGCGGTTTACGA 2066  
Db 1 GCAAGTGAACACAGAGTGTTCAAAGCGCTTCTGGGAGTGAAGGAGCGGTTTACGA 60  
Qy 2067 GTGACTTGGCTGAGCTCAGGGGCGGCACTGGACAGAACACACCTTGAAGCCAGCC 2126  
Db 61 GTGACTTGGCTGAGCTCAGGGGCGGCACTGGACAGAACACACCTTGAAGCCAGCC 120  
Qy 2127 TGGCTGGCCAGGCGAGCTGCTCTTCTCCGGGTTGTGACCCGCTCACTAGAGAG 2186  
Db 121 TGGCTGGCCAGGCGAGCTGCTCTTCTCCGGGTTGTGACCCGCTCACTAGAGAG 180  
Qy 2187 TTGGGGAAGCTTTTCACTTGGAGAGATTGCTCAACAACATGCTGGGCACTTGAACCT 2246  
Db 181 TTGGGGAAGCTTTTCACTTGGAGAGATTGCTCAACAACATGCTGGGCACTTGAACCT 240  
Qy 2247 CCTACTCTGCT 2258  
Db 241 CCTACTCTGCT 252

RESULT 10  
US-09-949-713-16  
Sequence 16, Application US/09949713  
Patent No. US20020044944A1  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, No. US20020044944A1  
APPLICANT: NAKAMURA, Shigekazu  
TITLE OF INVENTION: NOVEL FAS ANTIGEN DERIVATIVE  
FILE REFERENCE: 1110-207P  
CURRENT FILING DATE: US/09/949,713  
PRIOR FILING DATE: 2001-09-12  
PRIOR APPLICATION NUMBER: US/09/180,100  
PRIOR FILING DATE: 1998-11-02  
PRIOR APPLICATION NUMBER: PCT/JP97/01502  
PRIOR FILING DATE: 1997-05-01  
NUMBER OF SEQ ID NOS: 25

SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 16  
 LENGTH: 2534  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-949-713-16

Query Match  
 Best Local Similarity 7.0%; Score 226; DB 9; Length 2534;  
 Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2033 GAGCCTTCTGGGAGTGAAGGAGCGGTTTACGAGTACTTGGAGCTTCAGGGGCG 2092  
 DB 1 GAGCCTTCTGGGAGTGAAGGAGCGGTTTACGAGTACTTGGAGCTTCAGGGGCG 60  
 QY 2093 GGCAGTGGACGGAACACACCTTGAAGCCGCTGCTCCGAGGAGGCTGCTCTT 2152  
 DB 61 GGCAGTGGACGGAACACACCTTGAAGCCGCTGCTCCGAGGAGGCTGCTCTT 120  
 QY 2153 CTCGGGCGGTTGGTGAACCCGCTCACTACGAGTGGGAGAGCTCTTCACTTGGAGG 2212  
 DB 121 CTCGGGCGGTTGGTGAACCCGCTCACTACGAGTGGGAGAGCTCTTCACTTGGAGG 180  
 QY 2213 ATTGCTCAACACCATGTGCTGGCATCTGACCTCTTCACTTGGT 2258  
 DB 181 ATTGCTCAACACCATGTGCTGGCATCTGACCTCTTCACTTGGT 226

## RESULT 11

US-09-884-987-1  
 Sequence 1, Application US/09884987  
 Patent No. US20020102653A1  
 GENERAL INFORMATION:  
 APPLICANT: NAGATA, Shigekazu et al  
 TITLE OF INVENTION: DNA CODING FOR HUMAN CELL SURFACE ANTIGEN  
 FILE REFERENCE: 0020-4872P  
 CURRENT APPLICATION NUMBER: US/09/884,987  
 CURRENT FILING DATE: 2001-06-21  
 NUMBER OF SEQ ID NOS: 11  
 SOFTWARE: PatentIn version 3.0  
 SEQ ID NO 1  
 LENGTH: 2534  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: polyA\_site  
 LOCATION: (1831)..(1836)  
 NAME/KEY: mat\_peptide  
 LOCATION: (243)..()  
 NAME/KEY: sig\_peptide  
 LOCATION: (195)..(242)  
 NAME/KEY: CDS  
 LOCATION: (195)..(1199)  
 NAME/KEY: polyA\_site  
 LOCATION: (2352)..(2357)  
 NAME/KEY: polyA\_site  
 LOCATION: (2518)..(2532)  
 US-09-884-987-1

Query Match  
 Best Local Similarity 7.0%; Score 226; DB 10; Length 2534;  
 Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2033 GAGCCTTCTGGGAGTGAAGGAGCGGTTTACGAGTACTTGGAGCTTCAGGGGCG 2092  
 DB 1 GAGCCTTCTGGGAGTGAAGGAGCGGTTTACGAGTACTTGGAGCTTCAGGGGCG 60  
 QY 2093 GGCAGTGGACGGAACACACCTTGAAGCCGCTGCTCCGAGGAGGCTGCTCTT 2152  
 DB 61 GGCAGTGGACGGAACACACCTTGAAGCCGCTGCTCCGAGGAGGCTGCTCTT 120  
 QY 2153 CTCGGGCGGTTGGTGAACCCGCTCACTACGAGTGGGAGAGCTCTTCACTTGGAGG 2212

DB 121 CTCGGGCGGTTGGTGAACCCGCTCACTACGAGTGGGAGAGCTTTCACCTTGGAGG 180  
 QY 2213 ATTGCTCAACACCATGTGCTGGCATCTGACCTCTTCACTTGGT 2258  
 DB 181 ATTGCTCAACACCATGTGCTGGCATCTGACCTCTTCACTTGGT 226

## RESULT 12

US-10-027-632-145670/C  
 Sequence 145670, Application US/10027632  
 GENERAL INFORMATION:  
 APPLICANT: Wang, David G.  
 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
 FILE REFERENCE: 108827.129  
 CURRENT APPLICATION NUMBER: US/10/027,632  
 CURRENT FILING DATE: 2002-04-30  
 PRIOR APPLICATION NUMBER: US 60/218,006  
 PRIOR FILING DATE: 2000-07-12  
 PRIOR APPLICATION NUMBER: US 60/198,676  
 PRIOR FILING DATE: 2000-04-20  
 PRIOR APPLICATION NUMBER: US 60/193,483  
 PRIOR FILING DATE: 2000-03-29  
 PRIOR APPLICATION NUMBER: US 60/185,218  
 PRIOR FILING DATE: 2000-02-24  
 PRIOR APPLICATION NUMBER: US 60/167,363  
 PRIOR FILING DATE: 1999-11-23  
 PRIOR APPLICATION NUMBER: US 60/156,358  
 PRIOR FILING DATE: 1999-09-28  
 PRIOR APPLICATION NUMBER: US 60/146,002  
 PRIOR FILING DATE: 1999-08-09  
 NUMBER OF SEQ ID NOS: 325720  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 145670  
 LENGTH: 816  
 TYPE: DNA  
 ORGANISM: Human  
 US-10-027-632-145670

Query Match  
 Best Local Similarity 1.9%; Score 60.8; DB 13; Length 816;  
 Matches 83; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

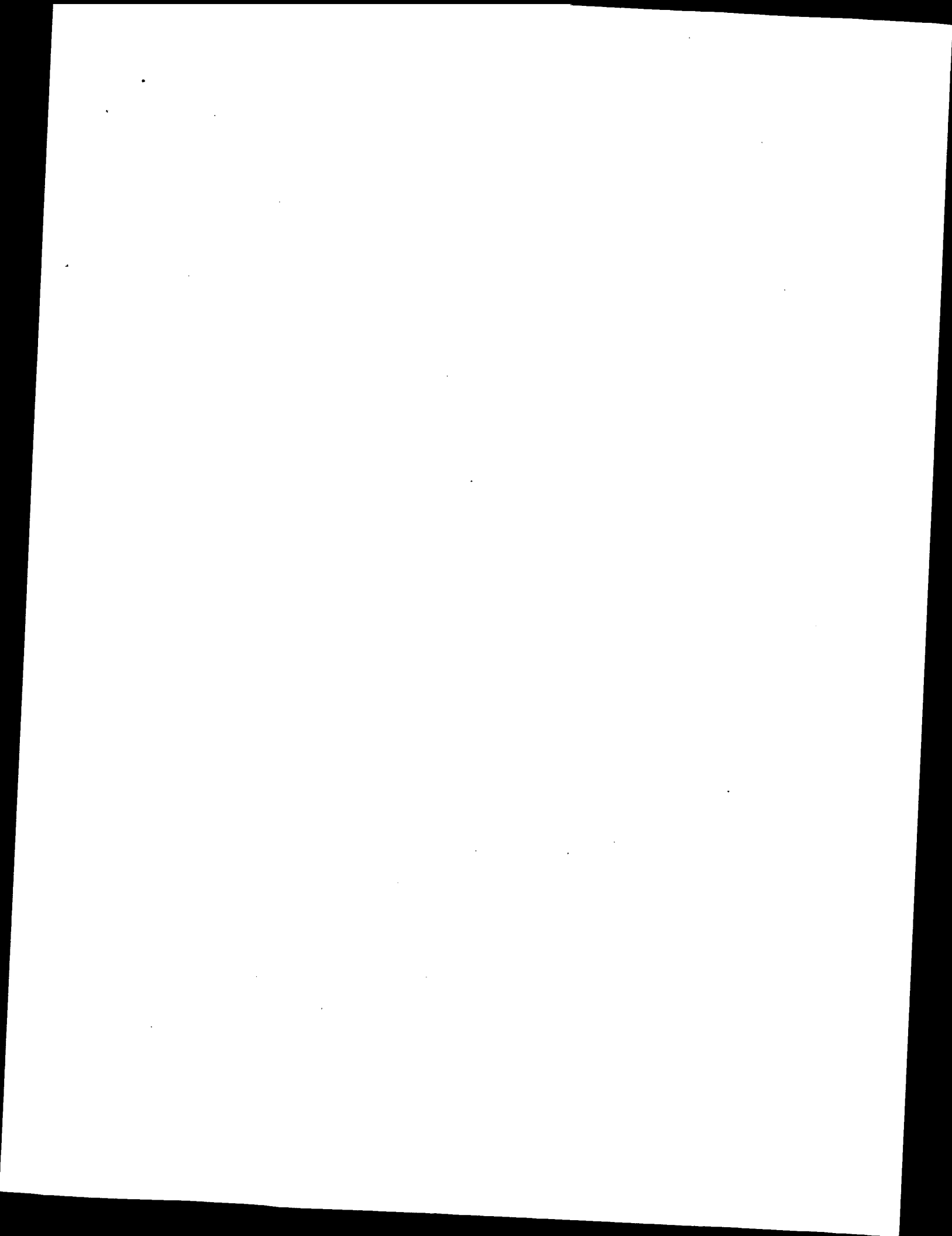
QY 697 TTTTATTTAAATGAACTTTTCACTTTGGAATAGTTTGAATTTCAAAAATTTGA 756  
 DB 307 TGTTTTGTTTAAATGAACTTTTCACTTTGGAATAGTTTGAATTTCAAAAATTTGA 248  
 QY 757 GAGATPATACAGAGATGCCATATACCATCTCTTATCCCACTTTCTTTGTGTAT 816  
 DB 247 AAGATPATACAGAAATATCCCATATCTGACACACCGTTTCCCTGTTTTACATTTAT 188

## RESULT 13

US-09-736-457-796/C  
 Sequence 796, Application US/09736457  
 Patent No. US2002016837A1  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Tongtong  
 APPLICANT: Bangur, Chaitanya S.  
 APPLICANT: Lodges, Michael A.  
 APPLICANT: Fanger, Gary  
 APPLICANT: Vedrick, Tom  
 APPLICANT: Carter, Darrick  
 APPLICANT: Retter, Marc  
 APPLICANT: Mannion, Jane  
 APPLICANT: Fan, Liqun  
 APPLICANT: Wang, Aijun  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 FILE REFERENCE: 210121.478C15  
 CURRENT APPLICATION NUMBER: US/09/736,457  
 CURRENT FILING DATE: 2000-12-13  
 NUMBER OF SEQ ID NOS: 1864

Page 11

1



Sun Aug 3 09:04:10 2003

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW nucleic - nucleic search, using sw model  
August 1, 2003, 08:37:27 ; Search time 78.1573 Seconds  
(without alignments)

Run on: 10468.541 Million cell updates/sec

US-09-834-291-6  
Title: 20  
Sequence: 1 ggaagagccctgacagcca 20

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmb1.\*  
1: gb\_da.\*  
2: gb\_hg.\*  
3: gb\_in.\*  
4: gb\_ov.\*  
5: gb\_ov.\*  
6: gb\_ov.\*  
7: gb\_ov.\*  
8: gb\_ov.\*  
9: gb\_ov.\*  
10: gb\_ov.\*  
11: gb\_ov.\*  
12: gb\_ov.\*  
13: gb\_ov.\*  
14: gb\_ov.\*  
15: gb\_ov.\*  
16: gb\_ov.\*  
17: gb\_ov.\*  
18: gb\_ov.\*  
19: gb\_ov.\*  
20: gb\_ov.\*  
21: gb\_ov.\*  
22: gb\_ov.\*  
23: gb\_ov.\*  
24: gb\_ov.\*  
25: gb\_ov.\*  
26: gb\_ov.\*  
27: gb\_ov.\*  
28: gb\_ov.\*  
29: gb\_ov.\*  
30: gb\_ov.\*  
31: gb\_ov.\*  
32: gb\_ov.\*  
33: gb\_ov.\*  
34: gb\_ov.\*  
35: gb\_ov.\*  
36: gb\_ov.\*  
37: gb\_ov.\*  
38: gb\_ov.\*  
39: gb\_ov.\*  
40: gb\_ov.\*  
41: gb\_ov.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query length	DB ID	Description
1	20	100.0	20	AX026094	AX026094 Sequence
2	18.4	92.0	20	AX026093	AX026093 Sequence
3	18.4	92.0	20	AX026095	AX026095 Sequence
4	18.4	92.0	20	AX026098	AX026098 Sequence
5	18.4	92.0	20	AX026112	AX026112 Sequence
6	18.4	92.0	266	AX026120	AX026120 Sequence
7	18.4	92.0	266	HSAD011034	AX026090 Sequence
8	18.4	92.0	280	AX026091	AX026091 Sequence
9	18.4	92.0	287	AX026092	AX026092 Sequence
10	18.4	92.0	3212	AX026089	AX026089 Sequence
11	18.4	92.0	45121	AX026085	AX026085 Sequence
12	18.4	92.0	187313	AX026089	AX026089 Sequence
13	18.4	92.0	167075	AX026089	AX026089 Sequence
14	18.4	92.0	167075	AX026089	AX026089 Sequence
15	17.4	87.0	111702	AX026089	AX026089 Sequence
16	17.4	87.0	150493	AX026089	AX026089 Sequence
17	17.4	87.0	154687	AX026089	AX026089 Sequence
18	17.4	87.0	191676	AX026089	AX026089 Sequence
19	17.4	87.0	224401	AX026089	AX026089 Sequence
20	17.4	87.0	224401	AX026089	AX026089 Sequence
21	17.4	87.0	224401	AX026089	AX026089 Sequence
22	16.8	84.0	402	AX026089	AX026089 Sequence
23	16.8	84.0	1208	AX026089	AX026089 Sequence
24	16.8	84.0	1263	AX026089	AX026089 Sequence
25	16.8	84.0	1923	AX026089	AX026089 Sequence
26	16.8	84.0	1923	AX026089	AX026089 Sequence
27	16.8	84.0	1923	AX026089	AX026089 Sequence
28	16.8	84.0	2461	AX026089	AX026089 Sequence
29	16.8	84.0	11459	AX026089	AX026089 Sequence
30	16.8	84.0	29518	AX026089	AX026089 Sequence
31	16.8	84.0	70408	AX026089	AX026089 Sequence
32	16.8	84.0	83120	AX026089	AX026089 Sequence
33	16.8	84.0	85040	AX026089	AX026089 Sequence
34	16.8	84.0	85040	AX026089	AX026089 Sequence
35	16.8	84.0	85040	AX026089	AX026089 Sequence
36	16.8	84.0	92431	AX026089	AX026089 Sequence
37	16.8	84.0	93431	AX026089	AX026089 Sequence
38	16.8	84.0	108336	AX026089	AX026089 Sequence
39	16.8	84.0	108336	AX026089	AX026089 Sequence
40	16.8	84.0	110000	AX026089	AX026089 Sequence
41	16.8	84.0	110000	AX026089	AX026089 Sequence
42	16.8	84.0	110000	AX026089	AX026089 Sequence
43	16.8	84.0	110382	AX026089	AX026089 Sequence
44	16.8	84.0	110382	AX026089	AX026089 Sequence
45	16.8	84.0	116090	AX026089	AX026089 Sequence

## ALIGNMENTS

RESULT 1	AX026094	20 bp	DNA	linear	PAF 16-SEP-2000
LOCUS	AX026094				
DEFINITION	Sequence 6 from Patent DE19847779.				
ACCESSION	AX026094				
VERSION	AX026094.1				
KEYWORDS	GI:10187525				
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE					
AUTHORS	Hueller-Schilling, M., Kramer, P. and Oren, M.				
TITLE	Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy				

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: DE 19847779-C 6 03-FEB-2000;  
DEUTSCHES KREBSFORSCH (DE)

## FEATURES

location/Qualifiers  
1..20

BASE COUNT /mol\_type="Homo sapiens"  
ORIGIN /db\_xref="taxon:9606"  
8 a 6 c 5 g 1 t

## Query Match

Best Local Similarity 100.0%; Score 20; DB 6; Length 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGAGCCCTGACAGCCA 20  
Db 1 GGAAGAGCCCTGACAGCCA 20

## RESULT 2

AX026093

LOCUS

DEFINITION Sequence 5 from Patent DE19847779.

ACCESSION AX026093

VERSION AX026093.1 GI:10187524

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

DEUTSCHES KREBSFORSCH (DE)

FEATURES

location/Qualifiers

1..20

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT

ORIGIN

7 a 7 c 5 g 1 t

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 6; Length 20;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAGAGCCCTGACAGCCA 20

Db 1 GGAAGAGCCCTGACAGCCA 20

## RESULT 3

AX026095

LOCUS

DEFINITION Sequence 7 from Patent DE19847779.

ACCESSION AX026095

VERSION AX026095.1 GI:10187526

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

DEUTSCHES KREBSFORSCH (DE)

FEATURES

location/Qualifiers

1..20

/organism="Homo sapiens"

BASE COUNT

ORIGIN

7 a 7 c 5 g 1 t



Query Match 92.0%; Score 18.4; DB 6; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 4.2e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAGAGCCCTGACAGCCA 20  
 1 GGACAGAGCCCTGACAGCCA 20

RESULT 6 AX026120 266 bp DNA linear PAT 16-SEP-2000

LOCUS AX026120 Sequence 32 from Patent DE19847779.

DEFINITION AX026120

AX026120.1 GI:10187551

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 Mueller-Schilling, M., Krammer, P. and Oren, M.

Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy

Patent: DE 19847779-C 32 03-FEB-2000;

DEUTSCHES KREBSFORSCH (DE)

location/Qualifiers

1. .266

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT 49 a 72 c 110 g 35 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 266;  
 Best Local Similarity 95.0%; Pred. No. 2.8e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAGAGCCCTGACAGCCA 20  
 160 GGACAGAGCCCTGACAGCCA 179

RESULT 7 HSA011034 266 bp DNA linear PRI 20-JAN-1999

LOCUS HSA011034 Homo sapiens DNA for enhancer of CD95 gene, partial.

DEFINITION AJ011034

VERSION AJ011034.1 GI:4165483

KEYWORDS CD95 gene; enhancer.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 Mueller, M., Wilder, S., Bannasch, D., Israeli, D., Lehnbach, K., Li-Weber, M., Friedman, S.L., Galle, P.R., Stremmel, W., Oren, M. and

Krammer, P.H. p53 activates the CD95 (APO-1/Fas) gene in response to DNA damage by anticancer drugs

J. Exp. Med. 188 (11), 2033-2045 (1998)

JOURNAL 99059827

DEFINITION 2 (bases 1 to 266)

KEYWORDS Direct Submision

Submitted (08-SEP-1998) Mueller M., Department of Internal Medicine

IV, University Hospital, Bergheimerstr. 58, Heidelberg, 69115,

GERMANY

location/Qualifiers

1. .266

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"  
 159. .178  
 /gene="CD95"  
 159. .178  
 /gene="CD95"  
 /evidence="experimental"

BASE COUNT 49 a 72 c 110 g 35 t

Query Match 92.0%; Score 18.4; DB 9; Length 266;  
 Best Local Similarity 95.0%; Pred. No. 2.8e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAGAGCCCTGACAGCCA 20  
 160 GGACAGAGCCCTGACAGCCA 179

RESULT 8 AX026090 720 bp DNA linear PAT 16-SEP-2000

LOCUS AX026090 Sequence 2 from Patent DE19847779.

DEFINITION AX026090

AX026090.1 GI:10187521

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 Mueller-Schilling, M., Krammer, P. and Oren, M.

Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy

Patent: DE 19847779-C 2 03-FEB-2000;

DEUTSCHES KREBSFORSCH (DE)

location/Qualifiers

1. .720

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT 142 a 181 c 216 g 181 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 720;  
 Best Local Similarity 95.0%; Pred. No. 2.4e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAGAGCCCTGACAGCCA 20  
 160 GGACAGAGCCCTGACAGCCA 179

RESULT 9 AX026091 2380 bp DNA linear PAT 16-SEP-2000

LOCUS AX026091 Sequence 3 from Patent DE19847779.

DEFINITION AX026091

VERSION AX026091.1 GI:10187522

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 Mueller-Schilling, M., Krammer, P. and Oren, M.

Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy

Patent: DE 19847779-C 3 03-FEB-2000;

DEUTSCHES KREBSFORSCH (DE)

location/Qualifiers

1. .2380

/organism="Homo sapiens"

/mol\_type="genomic DNA"

BASE COUNT 579 a /db\_xref="taxon:9606"  
 ORIGIN 595 c 568 g 638 t

Query Match 92.0%; Score 18.4; DB 6; Length 2380;  
 Best Local Similarity 95.0%; Pred. No. 2e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAAGCCCTGACAGCCA 20  
 DB 1820 GGACAGCCCTGACAGCCA 1839

RESULT 10  
 LOCUS AX026092 2827 bp DNA linear PAT 16-SEP-2000  
 DEFINITION Sequence 4 from Patent DE19847779.  
 ACCESSION AX026092  
 VERSION AX026092.1 GI:10187523  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.  
 TITLE Novel receptor dna useful for identifying apoptosis-modulating  
 JOURNAL substances potentially useful for cancer chemotherapy  
 DEUTSCHES KREBSFORSCH (DE)  
 PATENT: DE 19847779-C 4 03-FEB-2000;  
 FEATURES  
 source Location/Qualifiers  
 1..2827

BASE COUNT 728 a /organism="Homo sapiens"  
 ORIGIN /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606" 657 g 766 t

Query Match 92.0%; Score 18.4; DB 6; Length 2827;  
 Best Local Similarity 95.0%; Pred. No. 1.9e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAAGCCCTGACAGCCA 20  
 DB 2267 GGACAGCCCTGACAGCCA 2286

RESULT 11  
 LOCUS AX026089 3212 bp DNA linear PAT 16-SEP-2000  
 DEFINITION Sequence 1 from Patent DE19847779.  
 ACCESSION AX026089  
 VERSION AX026089.1 GI:10187520  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.  
 TITLE Novel receptor dna useful for identifying apoptosis-modulating  
 JOURNAL substances potentially useful for cancer chemotherapy  
 DEUTSCHES KREBSFORSCH (DE)  
 PATENT: DE 19847779-C 1 03-FEB-2000;  
 FEATURES  
 source Location/Qualifiers  
 1..3212

BASE COUNT 778 a /organism="Homo sapiens"  
 ORIGIN /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606" 809 g 841 t

Query Match 92.0%; Score 18.4; DB 6; Length 3212;

Best Local Similarity 95.0%; Pred. No. 1.9e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAAGCCCTGACAGCCA 20  
 DB 2659 GGACAGCCCTGACAGCCA 2678

RESULT 12  
 LOCUS AX695635 45121 bp DNA linear PAT 31-MAR-2003  
 DEFINITION Sequence 1262 from Patent WO03008583.  
 ACCESSION AX695635  
 VERSION AX695635.1 GI:29418787  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Morris, D.W. and Engelhard, E.K.  
 TITLE Novel compositions and methods for cancer  
 JOURNAL Patent: WO 03008583-A 1262 30-JAN-2003;  
 Sagres Discovery (US)  
 FEATURES  
 source Location/Qualifiers  
 1..45121  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606" 8836 c 9010 g 14049 t

BASE COUNT 13226 a /organism="Homo sapiens"  
 ORIGIN /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606" 8836 c 9010 g 14049 t

Query Match 92.0%; Score 18.4; DB 6; Length 45121;  
 Best Local Similarity 95.0%; Pred. No. 1.3e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAAGCCCTGACAGCCA 20  
 DB 10653 GGACAGCCCTGACAGCCA 10672

RESULT 13  
 LOCUS AL157394 187313 bp DNA linear PRI 22-AUG-2001  
 DEFINITION Human DNA sequence from clone Rpl1-399019 on chromosome 10,  
 complete sequence.  
 ACCESSION AL157394  
 VERSION AL157394.15 GI:15384622  
 KEYWORDS  
 SOURCE HTG.  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Blakey, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (22-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 On Aug 31, 2001 this sequence version replaced gi:14161146.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL, Sw, SWISSPROT, Tr, TrEMBL, Wp, WormPEP; Information on the WormPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr10>

RP11-399019 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pSpace.6  
This sequence is the entire insert of clone RP11-399019. The true left end of clone RP11-496H23 is at 166408 in this sequence. The true right end of clone RP11-30415 is at 18704 in this sequence.

## FEATURES

## SOURCE

1. 187313

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="10"

/clone\_1fb="RP11-399019"

/clone\_1fb="RP11-11.2"

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

/note="Sequence confirmed by AC015461 sequenced by WIBR."

## REFERENCE

Green, E.D.  
Submitted (05-FEB-2003) NIH Intramural Sequencing Center, 8717  
On Feb 5, 2003 this sequence version replaced gi:27884867.  
Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: <http://www.nisc.nih.gov>  
Contact: [nisc\\_zoo@nigri.nih.gov](mailto:nisc_zoo@nigri.nih.gov)  
Project Information  
Center project name: dvk  
Center clone name: 039j13

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

## Summary Statistics

Sequencing vector: plasmid; N/A; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 166451 bases at least Q40  
Consensus quality: 16589 bases at least Q30  
Consensus quality: 16648 bases at least Q20  
Insert size: 15000; agarose-fp  
Insert size: 166675; sum-of-contigs  
Quality coverage: 10.81x in Q20 bases; agarose-fp  
Quality coverage: 9.73x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

## FEATURES

## SOURCE

1. 167075

/organism="Didelphis virginiana"

/mol\_type="genomic DNA"

/db\_xref="taxon:9267"

/clone\_1fb="LB3-39J13"

/clone\_1fb="LB3"

/note="assembly\_fragment"

/note="assembly\_fragment"

/note="assembly\_fragment"

/note="assembly\_fragment"

/note="assembly\_fragment"

/note="assembly\_fragment"

RESULT 14  
AC139121  
LOCUS  
DEFINITION  
AC139121 167075 bp DNA linear HTG 05-FEB-2003  
Didelphis virginiana clone LB3-39J13, WORKING DRAFT SEQUENCE, 5  
ordered pieces.  
AC139121  
AC139121.2 GI:28209433  
HTG, HTGS\_PHASE2; HTGS\_DRAFT.  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (24-JAN-2003) NIH Intramural Sequencing Center, 8717

misc\_feature 133160.140947  
/note="assembly\_fragment"  
141048.167075  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:right"

BASE COUNT 48094 a 31698 c 33625 g 53258 t 400 others  
ORIGIN

Query Match 90.0%; Score 18; DB 2; Length 167075;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGACCCCTGACAGC 18  
DB 72852 GGAAGACCCCTGACAGC 72869

RESULT 15  
BV077964 714 bp DNA linear STS 31-MAY-2003  
LOCUS S212P60375FC6.T0 CZECHII/Ei Mus musculus STS genomic, sequence  
DEFINITION BV077964  
ACCESSION BV077964.1 GI:31193759  
KEYWORDS STS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 714)  
Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C.,  
Lander,B.S., Lindblad-Toh,K. and Daly,M.J.  
The mosaic structure of variation in the laboratory mouse genome  
Nature 420 (6915), 574-578 (2002)  
22354684  
PUBMED  
12466852  
COMMENT

Contact: Kerstin Lindblad-Toh  
Whitehead Institute for Biomedical Research, Center for Genome  
Research  
320 Charles Street, Cambridge, MA 02141, USA  
Tel: 6172521477  
Fax: 6172580903  
Email: kersli@genome.wi.mit.edu  
Primer A: None  
Primer B: None  
STS size: 714  
Protocol:

WGS-discovery: Paired-end low-coverage whole genome shotgun reads  
were generated from 129S1/SVimJ, C3H/HeJ, and BALB/cByJ. The WGS  
reads were placed uniquely on the MGSVC3 C57BL/6J assembly and SNP  
detection was carried out by SSAHA-SNP. 225,000 reads were  
annotated  
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J  
and the strain from which the particular read came. The validation  
rate for these SNPs was estimated at approximately 98%.  
location/Qualifiers

1. 714  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="CZECHII/Ei"  
/db\_xref="taxon:10090"  
/map="16 22-505 92651644-92652127"  
/clone\_lib="CZECHII/Ei"  
<1..>714

STS  
BASE COUNT 202 a 153 c 160 g 199 t  
ORIGIN

Query Match 87.0%; Score 17.4; DB 11; Length 714;  
Best Local Similarity 94.7%; Pred. No. 7.7e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGAAGACCCCTGACAGCCA 20  
DB 248 GGAAGACCCCTGACATGCCA 266

Search completed: August 1, 2003, 17:32:44  
Job time: 81.1573 secs

Sun Aug 3 09:04:15 2003

US-09-834-291-6.rst

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:13:22 ; Search time 68.4754 Seconds  
(without alignments)  
7098.748 Million cell updates/sec

Title: US-09-834-291-6

Perfect score: 1 ggaagaagccctgacacgca 20

Sequence: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Scoring table: 22781392 seqs, 12152238056 residues

Searched: 45562784

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthm:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_pbg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	18.4	92.0	793 12	BI763679 603049567
2	18.4	96.2	9	AL540709 AL540709
3	18.4	92.0	1089 13	BQ072834 AGENCOURT
4	17.4	87.0	255 13	BQ091744 UNM21B02

5	17.4	87.0	319 12	BI784680
6	17.4	87.0	693 10	BE272685
7	17.4	87.0	813 10	BE296494
8	17	85.0	1062 13	BQ070571
9	16.8	84.0	267 14	T95521
10	16.8	84.0	389 14	T79683
11	16.8	84.0	421 29	B2764825
12	16.8	84.0	426 9	AA233020
13	16.8	84.0	434 12	BM109280
14	16.8	84.0	440 13	BU962533
15	16.8	84.0	450 10	BE646301
16	16.8	84.0	454 14	T70290
17	16.8	84.0	455 10	BF175093
18	16.8	84.0	460 14	T87958
19	16.8	84.0	486 13	B0513664
20	16.8	84.0	490 28	BH028903
21	16.8	84.0	498 9	AV727532
22	16.8	84.0	508 28	BH029384
23	16.8	84.0	519 9	AI887588
24	16.8	84.0	521 10	BF962452
25	16.8	84.0	548 9	AA442710
26	16.8	84.0	554 9	AA490194
27	16.8	84.0	563 28	A2630360
28	16.8	84.0	569 13	BQ362540
29	16.8	84.0	577 12	B1924611
30	16.8	84.0	577 28	BH027622
31	16.8	84.0	587 10	AM950883
32	16.8	84.0	589 12	BM739855
33	16.8	84.0	590 9	AA625121
34	16.8	84.0	590 10	BE353461
35	16.8	84.0	591 13	BY470998
36	16.8	84.0	600 28	B2167939
37	16.8	84.0	605 13	BK474351
38	16.8	84.0	607 14	CB161114
39	16.8	84.0	642 28	BH800152
40	16.8	84.0	646 9	AV713997
41	16.8	84.0	659 9	AL8773982
42	16.8	84.0	669 13	BQ512401
43	16.8	84.0	672 13	BU266199
44	16.8	84.0	675 9	AL864641
45	16.8	84.0	677 12	BM700586

ALIGNMENTS

RESULT 1  
BI763679/c 793 bp mRNA linear EST 25-SEP-2001  
LOCUS 603049567F1 NIH\_MGC\_116 Homo sapiens CDNA clone IMAGE:5189752 5',  
DEFINITION mRNA sequence.  
ACCESSION BI763679  
VERSION BI763679.1 GI:15755257  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
1 (bases 1 to 793)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
CONTACT: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
CDNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L1AM11474 row: e column: 17  
High quality sequence stop: 786.

## FEATURES

source

Location/Qualifiers

1. 793

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5189752"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_116"

/note="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site: 1; NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH MGC Library."

## BASE COUNT

183 a 265 c 208 g 137 t

## Query Match

Best Local Similarity 92.0%; Score 18.4; DB 13; Length 793;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAAGCCCTGACAGCCA 20  
 Db 398 GGACAGCCCTGACAGCCA 379

## RESULT 2

AL540709/c

LOCUS

AL540709 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DB002YN18

ACCESSION

AL540709

VERSION

AL540709.2 GI:30544172

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS

L. W. B., Gruber, C., Jessee, J., and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

On Feb 15, 2001 this sequence version replaced gi:12871113.

COMMENT

Contract: Genoscope

Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

Library was

constructed by Life Technologies, a division of

Invitrogen. This

sequence belongs to sequence cluster 5554.r For

more information

about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DB002DG09QPI&amp;cluster=5554.r. Contact :

Feng Liang Email :

fengliang@life.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue

Genoscope sequence ID : CS0DB002DG09QPI.

Location/Qualifiers

1. 962

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DB002YN18"

/issue\_type="PLACENTA"

/clone\_lib="Homo sapiens PLACENTA"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed

with a NotI-oligo(dT) primer. Five prime and enriched,

the Not I and EcoRV sites of the pCMVSPORT 6 vector.

Library was not normalized."

## BASE COUNT

211 a 302 c 264 g 181 t 4 others

Query Match 92.0%; Score 18.4; DB 9; Length 962;  
 Best Local Similarity 95.0%; Pred. No. 2e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAAGCCCTGACAGCCA 20  
 Db 94 GGACAGCCCTGACAGCCA 75

## RESULT 3

BC072834

LOCUS

BC072834 AGNCOURT 6763280 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5756324

ACCESSION

BC072834

VERSION

BC072834.1 GI:19901880

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE

Unpublished

JOURNAL

Contact: Robert Strausberg, Ph.D.

COMMENT

Email: cgapbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)

found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LLM12796 row: 1 column: 21

High quality sequence stop: 645.

Location/Qualifiers

1. 1089

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5756324"

/issue\_type="Leukocyte"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_118"

/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV

(destroyed); RNA source leukocytes from anonymous pool of

non-activated adult donors. Library is oligo-dT primed

and directionally cloned (EcoRV site is destroyed upon

cloning). Average insert size 1.7 kb, insert size range

1.2-3.3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics tracking code 027. Note:

this is a NIH MGC Library."

BASE COUNT

280 a 263 c 333 g 212 t 1 others

ORIGIN

Location/Qualifiers

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 13; Length 1089;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAAGCCCTGACAGCCA 20  
 Db 369 GGACAGCCCTGACAGCCA 388

## RESULT 4

BC091744

LOCUS

BC091744 UMN21B02 Canine Brain CDNA Library Canis familiaris cDNA 5', mRNA

DEFINITION

255 bp mRNA linear EST 08-APR-2002



## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
1 (bases 1 to 693)  
AUTHORS  
NIH-MGC http://mgc.nci.nih.gov/  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: LICM39 row: j column: 15  
High quality sequence stop: 609.  
Location/Qualifiers  
1..693

## FEATURES

source

## BASE COUNT

197 a 137 c 212 g 147 t

## ORIGIN

Best Match  
Query Match  
Best Local Similarity 94.7%; Pred. No. 5.6e+02; Length 693;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 GAAAGCCCTGACAGCCA 20  
556 GAAAGCCCTGACAGCCA 574

## RESULT 7

BE296494  
LOCUS

## DEFINITION

601174770F1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:3530249 5',  
mRNA sequence.

## ACCESSION

BE296494

## VERSION

BE296494.1 GI:9180056

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

## REFERENCE

1 (bases 1 to 813)  
AUTHORS  
NIH-MGC http://mgc.nci.nih.gov/  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: LICM202 row: k column: 18  
High quality sequence stop: 536.  
Location/Qualifiers

## FEATURES

source

## source

1..813  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3530249"  
/tissue\_type="muscle"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: muscle; Vector: pOTB7, Site 1: XhoI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."  
BASE COUNT  
201 a 170 c 260 g 182 t

## ORIGIN

Best Match  
Query Match  
Best Local Similarity 94.7%; Pred. No. 5.9e+02; Length 813;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

## BASE COUNT

2 GAAAGCCCTGACAGCCA 20  
553 GAAAGCCCTGACAGCCA 571

## ORIGIN

Best Match  
Query Match  
Best Local Similarity 94.7%; Pred. No. 5.9e+02; Length 813;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

## RESULT 8

BE070571  
LOCUS

## DEFINITION

AGENCOURT\_683245 NIH\_MGC\_47 Homo sapiens cDNA clone IMAGE:5922971  
5', mRNA sequence.

## ACCESSION

BE070571

## VERSION

BE070571.1 GI:19899617

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

## REFERENCE

1 (bases 1 to 1062)  
AUTHORS  
NIH-MGC http://mgc.nci.nih.gov/  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LICM090 row: 1 column: 12  
High quality sequence stop: 624.  
Location/Qualifiers  
1..1062

## FEATURES

source

1..813  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5922971"  
/tissue\_type="muscle"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: muscle; Vector: pOTB7, Site 1: XhoI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."  
BASE COUNT  
201 a 170 c 260 g 182 t



BASE COUNT 197 a 352 c 317 g 196 t  
ORIGIN

Query Match 85.0%; Score 17; DB 13; Length 1062;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAG 17  
|||||  
768 GGAAGAGCCCTGACAG 764

RESULT 9 267 bp mRNA linear EST 27-MAR-1995  
T95521  
LOCUS Y646c1.1 Soares fetal liver spleen INFUS Homo sapiens cDNA clone  
DEFINITION IMAGE:120788 5' similar to gb:M55905 MALATE OXIDOREDUCTASE (HUMAN  
); mRNA sequence.  
ACCESSION T95521  
VERSION T95521.1 GI:734145  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman  
AUTHORS M., Holtman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,  
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaaskis, E., Waterston  
R., Williamson, A., Wohldmann, P. and Wilson, R.  
The Washu-Merck EST Project  
Unpublished  
TITLE Contract: Wilson RK  
JOURNAL Washington University School of Medicine  
COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Insert Size: 1976  
High quality sequence stops: 113 Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Insert Length: 1976 Std Error: 0.00  
Seq primer: M13RP1  
High quality sequence stop: 113.  
Location/Qualifiers  
1. 267  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:473333"  
/db\_xref="taxon:9606"  
/clone="IMAGE:120788"  
/sex="male"  
/dev stage="20 week-post conception fetus"  
/lab host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares fetal liver spleen INFUS"  
/note="Organ: Liver and Spleen; Vector: pTZ19 (Pharmacia)  
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5' AACTGAGAGATTAATTAAGATCTTTTCTTTTCTTTT 3']  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pTZ19 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M. Patricia Bonaldo."

BASE COUNT 72 a 54 c 57 g 76 t 8 others  
ORIGIN

Query Match 84.0%; Score 16.8; DB 14; Length 267;  
Best Local Similarity 90.0%; Pred. No. 7.6e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCA 20  
|||||

Db 136 GGAAGAGCCCTGACAGCA 155  
|||||

RESULT 10 389 bp mRNA linear EST 15-MAR-1995  
T79683  
LOCUS Y671c09.1 Soares fetal liver spleen INFUS Homo sapiens cDNA clone  
DEFINITION IMAGE:113680 5' similar to gb:M55905 MALATE OXIDOREDUCTASE (HUMAN  
); mRNA sequence.  
ACCESSION T79683  
VERSION T79683.1 GI:698192  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman  
AUTHORS M., Holtman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,  
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaaskis, E., Waterston  
R., Williamson, A., Wohldmann, P. and Wilson, R.  
The Washu-Merck EST Project  
Unpublished  
TITLE Other\_ESTs: y671c09.s1  
JOURNAL Washington University School of Medicine  
COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Insert Size: 1814  
High quality sequence stops: 244 Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Insert Length: 1814 Std Error: 0.00  
Seq primer: M13RP1  
High quality sequence stop: 244.  
Location/Qualifiers  
1. 389  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:469297"  
/db\_xref="taxon:9606"  
/clone="IMAGE:113680"  
/sex="male"  
/dev stage="20 week-post conception fetus"  
/lab host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares fetal liver spleen INFUS"  
/note="Organ: Liver and Spleen; Vector: pTZ19 (Pharmacia)  
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5' AACTGAGAGATTAATTAAGATCTTTTCTTTTCTTTT 3']  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pTZ19 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M. Patricia Bonaldo."

BASE COUNT 105 a 75 c 89 g 112 t 8 others  
ORIGIN

Query Match 84.0%; Score 16.8; DB 14; Length 389;  
Best Local Similarity 90.0%; Pred. No. 8.9e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCA 20  
|||||

Db 179 GGAAGAGCCCTGACAGCA 198  
|||||

RESULT 11 421 bp DNA linear GSS 13-MAR-2003  
B2764825/c  
LOCUS B2764825

DEFINITION SALK\_127083.30.15.n Arabidopsis thaliana TDNA insertion lines  
Arabidopsis thaliana genomic clone SALK\_127083.30.15.n, genomic  
survey sequence.

ACCESSION B2764825

VERSION B2764825

KEYWORDS GSS.

SOURCE GI:28937378

ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
1 (bases 1 to 421)

Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Gadgil, N.B.,  
C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, J., Shinn, P.,  
Zimmerman, J., and Ecker, J.R.

A Sequence-Indexed Library of Insertion Mutations in the  
Arabidopsis Genome

Unpublished

Contact: Joseph R. Ecker  
Salk Institute Genomic Analysis Laboratory (SIGAL)  
The Salk Institute for Biological Studies  
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
Tel: 858 453 4100 x1752  
Fax: 858 558 6379  
Email: ecker@salk.edu

This is single pass sequence recovered from the left border of  
TDNA. This sequence lies within an annotated intron of Atg24160.  
Class: TDNA tagged.

Location/Qualifiers

1..421

/organism="Arabidopsis thaliana"

/mol\_type="genomic DNA"

/strain="Columbia 0"

/db\_xref="taxon:3702"

/clone="SALK\_127083.30.15.n"

/note="PCR was performed on Arabidopsis thaliana lines  
each of which contains one or more TDNA insertion  
elements. The resultant fragment for each line was  
directly sequenced to determine the genomic sequence at  
the site of insertion. Details of the protocols used can  
be found at [http://sigal.salk.edu/tdna\\_protocols.html](http://sigal.salk.edu/tdna_protocols.html)"

BASE COUNT 113 a 77 c 87 g 144 t

ORIGIN

Query Match 84.0%; Score 16.8; DB 29; Length 421;  
Best Local Similarity 90.0%; Pred. No. 9.1e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCCA 20

DB 395 GAAAGAGCCCTGAAAGCCCA 376

RESULT 12

AA233020 426 bp mRNA linear EST 06-AUG-1997

LOCUS zrf4a08.r1 Soares NHMPU\_S1 Homo sapiens cDNA clone IMAGE:666422 5'

DEFINITION similar to gb:M55905 MALATE OXIDOREDUCTASE (HUMAN); mRNA sequence.

ACCESSION AA233020

VERSION AA233020

KEYWORDS EST.

SOURCE GI:1856013

ORGANISM Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 426)

Hallier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
Kudab, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, U., Moore, B.,  
Schallenberg, K., Steptoe, M., Tan, F., Theisling, B., White, Y., Wylie,  
T., Waterston, R., and Wilson, R.

WashU-Merck EST Project 1997

Unpublished

JOURNAL

TITLE

COMMENT

FEATURES

COMMENT

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)

This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.lnl.gov](mailto:info@image.lnl.gov)) for further information.  
Insert Length: 1352 Std Error: 0.00  
Seq primer: -26m3 rev2 EF from Amersham  
High quality sequence stop: 410.

Location/Qualifiers

1..426

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:5428454"

/db\_xref="taxon:9606"

/clone="IMAGE:666422"

/issue\_type="Pooled human melanocyte, fetal heart, and  
pregnant uterus"

/lab\_host="DH10B"

/clone\_lib="Soares NHMPU\_S1"

/note="Organ: mixed (see below); Vector: pT73D-Pac  
(Pharmacia) with a modified polylinker; Site 1: Not I;  
Site 2: Eco RI; Equal amounts of plasmid DNA from three  
normalized libraries (melanocyte ZMBW, pregnant uterus  
NBHPU, and fetal heart NBH19W) were mixed, and ss circles  
were made in vitro. Following HAP purification, this DNA  
was used as tracer in a subtractive hybridization  
reaction. The driver was PCR-amplified cDNAs from pools of  
5,000 clones made from the same 3 libraries. The pools  
consisted of I.M.A.G.E. clones 260232-265223,  
340488-345479, and 484488-489479."

BASE COUNT 119 a 90 c 96 g 121 t

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 426;  
Best Local Similarity 90.0%; Pred. No. 9.2e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCCA 20

DB 301 GCAAGAGCCCTGACAGCCCA 320

RESULT 13

BM109280 434 bp mRNA linear EST 10-MAR-2003

LOCUS BM109280/c

DEFINITION EST556816 potato roots Solanum tuberosum cDNA clone cPRO3B10 5' end

ACCESSION BM109280

VERSION BM109280

KEYWORDS EST.

SOURCE GI:17070044

ORGANISM Solanum tuberosum (potato)

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.

1 (bases 1 to 434)

van der Hoeven, R., Sun, H., Karayicheva, S.A., Teal, J., Van Aken, S.,  
Tanksley, S., and Baker, B.

Generation of ESTs from potato roots

Unpublished

Contact: Robin Buehl  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: [potato-array@tigr.org](mailto:potato-array@tigr.org)  
This clone can be obtained from the University of Arizona Genomics  
Institute. Orders can be made through URL:  
<http://genome.arizona.edu/orders/>  
Seq primer: T3.

Location/Qualifiers

FEATURES

source

1. 434

/organism="Solanum tuberosum"

/mol\_type="mRNA"

/cultivar="Kennebec"

/db\_xref="taxon:4113"

/clone="CPR03B10"

/tissue\_type="roots"

/dev\_stage="in vitro grown stem cuttings"

/lab\_host="SOLR"

/clone\_1lb="potato roots"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University, Tanksley lab; sequencing: The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after planting the stem cuttings from in vitro grown plants on medium."

BASE COUNT 125 a 98 c 82 g 129 t

ORIGIN

Query Match 84.0%; Score 16.8; DB 12; Length 434;  
Best Local Similarity 90.0%; Pred. No. 9.2e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCA 20  
|||||  
405 GGAAGAGCCCTGACAGCA 386

Db

RESULT 14  
BU962533 440 bp mRNA linear EST 21-OCT-2002  
LOCUS  
AGENCOURT 10616779 NIH\_MGC 169 Mus musculus cDNA clone  
IMAGE:6743287 5', mRNA sequence.  
BU962533  
BU962533.1 GI:24192105  
EST.  
Mus musculus (house mouse)  
SOURCE  
Mus musculus  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 440)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Dr. Jonathan Kuo, NIMH  
CDNA Library Preparation: Michael Brownstein Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov 1 column: 06  
Plate: LMC3082 row: 1  
High quality sequence stop: 440.  
Location/Qualifiers  
1. 440  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6743287"  
/lab\_host="DH10B (T1-phage-resistant)"  
/clone\_1lb="NIH\_MGC 169"  
/note="Organ: Testicles; Vector: pDNR-LIB; Site 1: SfiI  
(ggcattatggcc); Site 2: SfiI (ggcgcgcgcgc); cDNA made  
by oligo-dT priming and directionally cloned. 5' and 3'  
adaptors were used in cloning as follows:  
5'-AAGCAGTGTATCAACGACAGTGGCATTCAGCGCGG-3' and  
5'-ATTCTAGAGCCGAGCGCGCGCATATG-dT(30)NN-3'. Full-length  
enriched library was constructed using the Clontech  
Creator SMART kit and size-selected to contain the 0.5 kb  
size fraction. Library created in the laboratory of M.  
Brownstein (NIMH, NIH). Note: this is a NIH\_MGC library."

FEATURES  
source

BASE COUNT 188 a 81 c 74 g 95 t 2 others

ORIGIN

Query Match 84.0%; Score 16.8; DB 13; Length 440;  
Best Local Similarity 90.0%; Pred. No. 9.3e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCA 20  
|||||  
8 GGAAGAGCCCTGACAGCA 27

Db

RESULT 15  
BE646301 450 bp mRNA linear EST 05-SEP-2000  
LOCUS  
7864d08.x1 NCI\_CGAP\_P128 Homo sapiens cDNA clone IMAGE:3289167 3',  
mRNA sequence.  
ACCESSION  
BE646301 GI:9970612  
VERSION  
BE646301.1  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 450)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL, send email to:  
info@image.llnl.gov  
Seq primer: -40UP from Gbco.  
Location/Qualifiers  
1. 450  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3289167"  
/sex="male"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_1lb="NCI\_CGAP\_P128"  
/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia)  
with a modified polylinker; Plasmid DNA from the  
normalized library NCI CGAP\_P128 was prepared, and ss  
circles were made in vitro. Following HAP purification,  
this DNA was used as tracer in a subtractive hybridization  
reaction. The driver was PCR-amplified cDNAs from a pool  
of 5,000 clones made from the same library (cloneids  
985608-986759, 1101192-1101959, and 1217928-1220615).  
Subtraction by Bento Soares and M. Fatima Bonardo."

BASE COUNT 87 a 135 c 116 g 112 t

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 450;  
Best Local Similarity 90.0%; Pred. No. 9.4e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCA 20  
|||||  
433 GGAAGAGCCCTGACAGCA 414

Db

Search completed: August 1, 2003, 20:42:54  
Job time : 72.4754 secs

Sun Aug 3 09:04:15 2003

us-09-834-291-6.rst

Page 8

Sun Aug 3 09:04:11 2003

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:32:17 ; Search time 6.69868 Seconds  
(without alignments)  
8059.612 Million cell updates/sec

Title: US-09-834-291-6

Perfect score: 1 ggaagagccctgacagca 20

Sequence:

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

5105512

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_190un03.\*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.4	92.0	266	21	AAZ88700 Human CP95 recepto
2	16.8	84.0	402	25	AAZ88700 Novel rti cDNA can
3	16.8	84.0	502	25	AAZ88700 cDNA encoding nove
4	16.8	84.0	1444	13	AAO22658 Sequence encoding
5	16.8	84.0	1752	13	AAO23258 Mitochondrial NAD(
6	16.8	84.0	1923	24	ABL67216 Thyroid cancer rel
7	16.8	84.0	1963	24	ABL67216 Human polynucleoti
8	16.8	84.0	1963	24	ABL67216 Human polynucleoti

US-09-834-291-6.rng

9	16.8	84.0	2390	22	AAZ88700
10	16.8	84.0	2463	22	AAZ88700
11	16.8	84.0	3085	22	AAZ88700
12	16.8	84.0	3085	22	AAZ88700
13	16.4	82.0	936	24	AAZ88700
14	16.4	82.0	2214	21	AAZ88700
15	16.4	82.0	2214	25	AAZ88700
16	16.4	82.0	2214	25	AAZ88700
17	15.8	79.0	251	24	AAZ88700
18	15.8	79.0	345	21	AAZ88700
19	15.8	79.0	500	24	AAZ88700
20	15.8	79.0	624	23	AAZ88700
21	15.8	79.0	4057	17	AAZ88700
22	15.8	79.0	4796	22	AAZ88700
23	15.8	79.0	4796	22	AAZ88700
24	15.8	79.0	10519	23	AAZ88700
25	15.8	79.0	10519	23	AAZ88700
26	15.8	79.0	10519	23	AAZ88700
27	15.8	79.0	13337	22	AAZ88700
28	15.8	79.0	13337	23	AAZ88700
29	15.8	79.0	13337	23	AAZ88700
30	15.8	79.0	21129	22	AAZ88700
31	15.8	79.0	22700	22	AAZ88700
32	15.8	79.0	29255	23	AAZ88700
33	15.8	79.0	31813	22	AAZ88700
34	15.8	79.0	31813	22	AAZ88700
35	15.8	79.0	31813	22	AAZ88700
36	15.8	79.0	110608	24	AAZ88700
37	15.8	79.0	349980	24	AAZ88700
38	15.4	77.0	508	21	AAZ88700
39	15.4	77.0	508	22	AAZ88700
40	15.4	77.0	508	25	AAZ88700
41	15.4	77.0	1253	24	AAZ88700
42	15.4	77.0	3812	10	AAZ88700
43	15.4	77.0	465237	24	AAZ88700
44	15.4	77.0	465237	24	AAZ88700
45	15.4	77.0	465237	25	AAZ88700

## ALIGNMENTS

RESULT 1  
AAZ88700 standard: DNA; 266 BP.

AAZ88700	AAZ88700	standard: DNA; 266 BP.
XX	XX	
AC	AAZ88700	
XX	XX	
DT	11-MAY-2000	(first entry)
XX	XX	
DE	Human CP95 receptor intron 1 fragment.	
XX	XX	
KW	p53; CD95 receptor; human; screening; apoptosis-modulation;	
KM	cancer chemotherapy; ss.	
XX	XX	
OS	Homo sapiens.	
XX	XX	
FT	Key	Location/Qualifiers
FT	protein_bind	160..179
FT		/tag= a
FT		/bound_molety= p53
XX	XX	
PN	DE19847779-CL.	
PD	03-FEB-2000.	
XX	XX	
PF	16-OCT-1998;	98DE-1047779.
XX	XX	
PR	16-OCT-1998;	98DE-1047779.
XX	XX	
PA	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.	
XX	XX	
PI	Krammer P, Mueller-Schilling M, Oren M;	

CDNA encoding nove  
Novel protein kina  
Lactococcal promot  
Streptococcus poly  
Human secreted pro  
Human secreted pro  
Secreted protein-e  
Receptor #71 parti  
Human ORX polynu  
Human colon cancer  
Human genome-deriv  
Human prostate exp  
Mouse neuron syst  
Human nervous syst  
Genomic sequence #  
Genomic sequence #  
Human nervous syst  
Genomic sequence #  
Genomic sequence #  
Human DNA-binding  
Human cardiovascular  
Human immune/haema  
Proionibacterium  
Human ovarian and  
Human reproductive  
Human CDNA differe  
Bifidobacterium lo  
CDNA encoding huma  
Colon tumour relat  
Human lung specif  
Human lung specif  
Human oestrogen re  
Gene encoding huma

XX WPI: 2000-162245/15.  
XX

PT Novel receptor DNA useful for identifying apoptosis-modulating  
XX substances potentially useful for cancer chemotherapy  
XX

PS Claim 2; Fig 4; 12pp; German.

CC This invention describes a novel p53-binding region of a human CD95  
CC receptor DNA molecule. The p53-binding region, or a vector containing  
CC it, can be used to screen for apoptosis-modulating substances  
CC potentially useful for cancer chemotherapy. This sequence represents a  
CC fragment of the human CD95 receptor intron 1 which contains a p53 binding  
CC region described in the method of the invention.  
XX

SO Sequence 266 BP; 49 A; 72 C; 110 G; 35 T; 0 other;

Query Match  
Best Local Similarity 92.0%; Score 18.4; DB 21; Length 266;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCCA 20  
DB 160 GGAAGAGCCCTGACAGCCCA 179

RESULT 2  
ID AAF64773 standard; cDNA; 402 BP.

AC AAF64773;  
XX

DT 09-APR-2001 (first entry)

XX Novel human polynucleotide, SEQ ID NO: 529.

XX Human; cyclostatic; gene therapy; colon cancer; prostate cancer;  
XX breast cancer; lung cancer; cancer detection; ss.

OS Homo sapiens.

XX W0200102568-A2.

PD 11-JAN-2001.

PF 30-JUN-2000; 2000WO-US18374.

PR 02-JUL-1999; 99US-0142310.

XX 02-JUL-1999; 99US-0142311.

PA (CHIR) CHIRON CORP.

XX (HYSE-) HYSEQ INC.

PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;  
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson J, Kassam A;  
PI Cirenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;  
PI Kita D, Garcia V, Jones LW, Strache-Crain B;  
XX WPI: 2001-091805/10.

PT Library of polynucleotides for diagnosing a cancerous state of a  
PT mammalian cell and detecting cancer, particularly of the colon or  
PT prostate, comprises 3351 human polynucleotide sequences -  
XX Claim 9; Page 620; 1046pp; English.

CC The present sequence is one of 3351 sequences in a library of human  
CC polynucleotides. The library is used to detect differentially expressed  
CC genes correlated with a cancerous state of a mammalian cell and can  
CC detect colon, prostate, breast and lung cancer. The library can be used  
CC to produce probes for detection of mRNA and to produce additional copies  
CC of the polynucleotides. The probes can be used for chromosome mapping of  
CC the polynucleotide and for detection of transcription levels. Ribozymes

CC or antisense oligonucleotides can be generated. The polynucleotides and  
CC their gene products are used as genetic or biochemical markers (e.g. in  
CC blood or tissues) that will detect the earliest changes along the  
CC carcinogenesis pathway and/or monitor the efficacy of therapies and  
CC preventive interventions. The polynucleotides, polypeptides and  
CC antibodies against them can be used in pharmaceutical compositions to  
CC treat the cancers and proliferative disorders such as neoplasia,  
CC dysplasia and hyperplasia.  
XX

SO Sequence 402 BP; 72 A; 121 C; 102 G; 106 T; 1 other;

Query Match  
Best Local Similarity 84.0%; Score 16.8; DB 22; Length 402;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCCA 20  
DB 346 GGAAGAGCCCTGACAGCCCA 365

RESULT 3

ID ABZ19870 standard; cDNA; 502 BP.

XX ABZ19870;

DT 23-JAN-2003 (first entry)

XX Group III cDNA cancer related clone SEQ ID NO:2296.

XX Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;  
XX immune response; virology; immunology; microbiology; molecular biology;  
XX recombinant DNA technology; gene; ss.

OS Homo sapiens.

XX W0200278516-A2.

PD 10-OCT-2002.

PF 28-MAR-2002; 2002WO-US10421.

PR 30-MAR-2001; 2001US-280255P.

XX 28-AUG-2001; 2001US-31563P.

XX 09-JAN-2002; 2002US-347313P.

PA (CORI-) CORIXA CORP.

XX Wang T, Wang S, Bangur CS, Gaiger A;

XX WPI: 2003-058387/05.

PT New immunogenic polynucleotides or polypeptides useful for diagnosing,  
PT preventing and treating cancer expressing CT or CP mRNA antigens, and  
PT recombinant DNA techniques -  
XX Claim 1; SEQ ID 2296; 207pp; English.

CC ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and  
CC ABP54446 to ABP54472 represent protein (II) sequences, from the present  
CC invention. (I) and (II) have cytostatic activity and can be used in gene  
CC therapy and vaccines. (I), (II), antibodies and compositions from the  
CC present invention are useful for diagnosing, preventing and treating  
CC cancer, which expresses CT or CP mRNA antigens. They are useful for  
CC immunology, microbiology, molecular biology and recombinant DNA  
CC techniques.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

SO Sequence 502 BP; 131 A; 124 C; 163 G; 84 T; 0 other;

Query Match 84.0%; Score 16.8; DB 25; Length 502;  
Beet local similarity 90.0%; Pred. No. 87;  
Matches 18; Conservative 2; Indels 0; Gaps 0;  
1 GGAAAGCCCTGACAGCCA 20  
159 GGCAAGCCCTCACAAGCCA 178

RESULT 4  
AAS40992 standard; cDNA; 1444 BP.  
ID AAS40992:  
AC AAS40992:  
XX 17-DEC-2001 (first entry)  
DT CDNA encoding novel human enzyme polypeptide #208.  
XX  
DE Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;  
XX ligase; hyperproliferative disorder; immunodeficiency disorder;  
XX autoimmune disorder; neurological disorder; metabolic disorder;  
XX inflammatory disorder; cardiovascular disorder; reproductive disorder;  
XX blood-related disorder; infectious disorder; gene therapy; cytostatic;  
XX anti arthritic; nephrotoxic; anticoagulant; ss.  
XX Homo sapiens.  
OS WO200155301-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01239.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0214886.  
XX 28-JUN-2000; 2000US-0215335.  
XX 30-JUN-2000; 2000US-0216860.  
XX 07-JUL-2000; 2000US-0216860.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0218290.  
XX 14-JUL-2000; 2000US-0220963.  
XX 26-JUL-2000; 2000US-0220964.  
XX 14-AUG-2000; 2000US-0224518.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225267.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225477.  
XX 14-AUG-2000; 2000US-0225757.  
XX 14-AUG-2000; 2000US-0225759.  
XX 14-AUG-2000; 2000US-0226681.  
XX 18-AUG-2000; 2000US-0226827.  
XX 22-AUG-2000; 2000US-0227182.  
XX 22-AUG-2000; 2000US-0227182.  
XX 23-AUG-2000; 2000US-0228224.  
XX 30-AUG-2000; 2000US-0228287.  
XX 01-SEP-2000; 2000US-0229343.  
XX 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231245.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0231415.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0232403.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235634.  
PR 27-SEP-2000; 2000US-0235636.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 29-SEP-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240360.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.

PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 01-DEC-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 05-DEC-2000; 2000US-0250392.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 06-DEC-2000; 2000US-0256719.  
 PR 08-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 11-DEC-2000; 2000US-0251989.  
 PR 05-JUN-2001; 2000US-0254097.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 FI Rosen CA, Barash SC, Ruben SM;  
 DR WPI; 2001-465566/50.  
 XX P-PSDB; AAU23122.

Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases -

Claim 4; SEQ ID No 218; 1180bp; English.

The present invention relates to the isolation of novel human enzyme polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of disorders including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. influenza). The polynucleotides of the invention can also be used in gene therapy. AAG40785-AAG41684 represent cDNA sequences encoding for the novel enzyme polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 1444 BP; 432 A; 284 C; 325 G; 403 T; 0 other;

Query Match  
 Best Local Similarity 84.0%; Score 16.8; DB 22; Length 1444;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GGAAGAGCCCTGACAGCA 20  
 DB 960 GCAAGAGCCCTGACAGCA 979

RESULT 5  
 AAQ22695

ID AAQ22695 standard; DNA; 1752 BP.  
 XX AC AAQ22695;  
 XX AC  
 DT 25-MAR-2003 (updated)  
 DT 24-JUL-1992 (first entry)  
 XX DE  
 XX DE Sequence encoding mitochondrial NAD(P)+-dependent malate enzyme.  
 KW Carbon metabolism; pyruvate formation; ss.  
 OS Homo sapiens.  
 XX OS  
 XX OS  
 FT Key  
 FT CDS Location/Qualifiers  
 FT 1..1752  
 FT /\*tag= a  
 XX PN  
 XX PD DE4028618-A.  
 XX PD 12-MAR-1992.  
 XX PF  
 XX PF 08-SEP-1990; 90DE-4028618.  
 XX PR 08-SEP-1990; 90DE-4028618.  
 XX PR 19-JUN-1991; 91DE-4120178.  
 XX PA (BOEH) BOEHRINGER INGELHEIM.  
 XX PI Dworkin MB, Leob, Leober G, Krystek E, Maurerfoggy I;  
 XX DR WPI; 1992-069407/12.  
 XX DR P-PSDB; AAR21845.  
 XX PT Human mitochondrial NAD(P)-dependent malate enzyme - used to  
 XX PT study formation of pyruvate from amino acid(s) in tumour cells  
 XX PS  
 XX PS Claim 2; Page 12-13; 20bp; German.  
 CC The inventors claim mitochondrial NAD(P)+-dependent malate enzyme  
 CC and DNA encoding it. AAR21845 has 5' and 3' non-coding regions.  
 CC The enzyme catalyses conversion of malate to pyruvate. Both the DNA and  
 CC the enzyme are useful for studying carbon metabolism in rapidly  
 CC dividing cells, esp. pyruvate formation from amino acids.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX CC  
 XX CC  
 SQ Sequence 1752 BP; 562 A; 326 C; 386 G; 478 T; 0 other;  
 Query Match  
 Best Local Similarity 84.0%; Score 16.8; DB 13; Length 1752;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GGAAGAGCCCTGACAGCA 20  
 DB 1483 GCAAGAGCCCTGACAGCA 1502

RESULT 6

ID AAQ23258  
 ID AAQ23258 standard; DNA; 1923 BP.  
 XX AC AAQ23258;  
 XX AC  
 DT 31-JUL-1992 (first entry)  
 DT 31-JUL-1992 (first entry)  
 XX DE  
 XX DE Mitochondrial NAD(P)+-dependent malate enzyme.  
 KW C-metabolism; tumour; pyruvate; T-lymphocyte; ss.  
 OS Homo sapiens.  
 XX OS  
 XX OS  
 FT Key  
 FT CDS Location/Qualifiers  
 FT 90..1844



Sun Aug 3 09:04:11 2003

us-09-834-291-6.rng

```

FT      /tag= a
FT      /product= malate_enzyme
FT      90..149
FT      sig.peptide
FT      /tag= b
FT      150..1844
FT      mat.peptide
FT      /tag= c
FT      1905..1910
FT      polyA_signal
FT      /tag= d
FT      /note= "homologous to poly(A) signal AATAAA"
XX      WC9204448-A.
XX      19-MAR-1992.
XX      23-AUG-1991; 91MO-EP01602.
XX      19-JUN-1991; 91DE-4120178.
XX      08-SEP-1990; 90DE-4028618.
XX      (BOEH ) BOEHRINGER INGELHEIM.
XX      Dworkin MB, Loeber G, Krystek E, Maurelfof I, Frubels B;
XX      WPI: 1992-114355/14.
XX      P-PSDB; AAR23356.
XX      New human mitochondrial malate enzyme and DNA encoding it - for
XX      studying carbon metabolism in cells, also specific antibodies for
XX      purification and assay
XX      Claim 1; Page 46 + Fig 3; 60pp; German.
XX      The sequence may be used to study C-metabolism in rapidly dividing
XX      (tumor) cells, esp. pyruvate formation from amino acids
XX      The enzyme was first isolated from the supernatant of mitochondrial
XX      preparations from the transformed human T-lymphocyte cell line
XX      1301. Tryptic fragments were partially sequenced and used as a
XX      basis for the design of oligonucleotides. These were used in PCR
XX      for amplification of malate enzyme encoding DNA in a cDNA bank
XX      CC prep. from fibrosarcoma Hs913. Amplified fragments were subcloned
XX      in pUC18, sequenced and used to probe the fibrosarcoma bank.
XX      A 1923bp insert was isolated and cloned in Bluescript KS+.
XX      The poly(A+) tail is not included in this sequence.
XX      Sequence 1923 BP; 598 A; 373 C; 436 G; 516 T; 0 other;
XX      Query Match 84.0%; Score 16.8; DB 13; Length 1923;
XX      Best Local Similarity 90.0%; Pred. No 1e+02; Indels 0; Gaps 0;
XX      Matches 18; Conservative 0; Mismatches 2;
XX      1 GGAAGAGCCCTGACAGCA 20
XX      1572 GCAAGGCCCCGACAGCA 1591
XX      Db
XX      RESULT 7
XX      ABL67216
XX      ID ABL67216 standard; DNA; 1923 BP.
XX      AC ABL67216;
XX      DT 15-MAY-2002 (first entry)
XX      DE Thyroid cancer related gene sequence SEQ ID NO:5553.
XX      KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX      stomach; lung; prostate; carcinoma; antitumor; cancerous;
XX      cytostatic; gene therapy; antineoplastic; Wilms tumor; adenocarcinoma;
XX      gene; ds.
XX      KW Homo sapiens.
XX      OS
XX      PN WO200194629-A2.

```

```

XX      13-DEC-2001.
XX      30-MAY-2001; 2001MO-US10838.
XX      05-JUN-2000; 2000US-209473P.
XX      05-JUN-2000; 2000US-209531P.
XX      18-SEP-2000; 2000US-233133P.
XX      18-SEP-2000; 2000US-23317P.
XX      20-SEP-2000; 2000US-234009P.
XX      20-SEP-2000; 2000US-234034P.
XX      20-SEP-2000; 2000US-234052P.
XX      20-SEP-2000; 2000US-234509P.
XX      22-SEP-2000; 2000US-234567P.
XX      22-SEP-2000; 2000US-234923P.
XX      25-SEP-2000; 2000US-235077P.
XX      25-SEP-2000; 2000US-235083P.
XX      25-SEP-2000; 2000US-235134P.
XX      25-SEP-2000; 2000US-235280P.
XX      26-SEP-2000; 2000US-235639P.
XX      26-SEP-2000; 2000US-235711P.
XX      27-SEP-2000; 2000US-235720P.
XX      27-SEP-2000; 2000US-235840P.
XX      27-SEP-2000; 2000US-235842P.
XX      28-SEP-2000; 2000US-236032P.
XX      28-SEP-2000; 2000US-236034P.
XX      28-SEP-2000; 2000US-236109P.
XX      28-SEP-2000; 2000US-236111P.
XX      28-SEP-2000; 2000US-236842P.
XX      29-SEP-2000; 2000US-236891P.
XX      29-SEP-2000; 2000US-237172P.
XX      02-OCT-2000; 2000US-237173P.
XX      02-OCT-2000; 2000US-237294P.
XX      02-OCT-2000; 2000US-237295P.
XX      02-OCT-2000; 2000US-237316P.
XX      02-OCT-2000; 2000US-237425P.
XX      03-OCT-2000; 2000US-237588P.
XX      03-OCT-2000; 2000US-237604P.
XX      03-OCT-2000; 2000US-237606P.
XX      03-OCT-2000; 2000US-237608P.
XX      01-NOV-2000; 2000US-244867P.
XX      01-NOV-2000; 2000US-245084P.
XX      (AVAIL-) AVALON PHARM.
XX      Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX      PI Soppet DR, Weaver Z;
XX      WPI: 2002-188264/24.
XX      Screening for anti-neoplastic agent involves exposing cells to a
XX      chemical agent to be tested for anti-neoplastic activity, and
XX      determining a change in expression of a gene of a signature gene set
XX      Claim 1; SEQ ID 5553; 44pp; English.
XX      The present invention describes a method (M1) for screening for an
XX      anti-neoplastic agent. The method involves exposing cells to a chemical
XX      agent to be tested for anti-neoplastic activity, determining a change in
XX      expression of at least one gene (I) of a signature gene set where (I)
XX      comprises a sequence (S) selected from 8447 sequences (given in ABL6164
XX      CC ABL6164) or is at least 95% identical to (S), where a change in
XX      expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX      activity and can be used in gene therapy. M1 can be used for screening
XX      an anti-neoplastic agent, and can be used for producing a product which
XX      is the data collected with respect to the anti-neoplastic agent as a
XX      result of M1, and the data is sufficient to convey the chemical
XX      structure and/or properties of the agent. M1 can be used in the

```

CC Treatment of cancer such as colon, breast, stomach, lung, thyroid,  
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
 CC carcinoma, papillary carcinoma and Wilms' tumour.  
 XX Sequence 1923 BP; 599 A; 373 C; 436 G; 515 T; 0 other.

Query Match 84.0%; Score 16.8; DB 24; Length 1923;  
 Best Local Similarity 90.0%; Pred. No. 1e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GGAAGAGCCCTGACAGCCA 20  
 Db 1572 GCAAGGCCCTGACAGCCA 1591

RESULT 8  
 ABL89871  
 ID ABL89871 standard; cDNA; 1963 BP.  
 XX  
 AC ABL89871;  
 XX  
 DT 24-MAY-2002 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 433.

KW Cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral;  
 KW anti-allergic; hepatotropic; antidiabetic; antihistaminic; antifungal;  
 KW vulnery; anticonvulsant; antibacterial; antiparasitic; antitumor;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular;  
 KW neurological disease; infection; human; secreted protein; gene; ss.  
 OS Homo sapiens.

XX WO200190304-A2.  
 XX  
 PD 29-NOV-2001.

PF 18-MAY-2001; 2001WO-US16450.

PR 19-MAY-2000; 2000US-205515P.

(HUMA-) HUMAN GENOME SCI INC.

Birse CE, Rosen CA;

WPI; 2002-122018/16.

P-PSDB; ABB89462.

Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
 prevention of neutral, immune system, muscular, reproductive, and  
 gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
 disorders.

Claim 4, SEQ ID NO 433; 2081bp + Sequence Listing; English.

The invention relates to novel genes (ABL89449-ABL90853) and proteins  
 (ABB89040-ABB90444) useful for preventing, treating or ameliorating  
 medical conditions e.g. by protein or gene therapy. The genes are  
 isolated from a range of human tissues disclosed in the specification.  
 The nucleic acids, proteins, antibodies and (ant)agonists are useful  
 in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
 and ovarian cancer and other cancers of the adrenal gland, bone, bone  
 marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
 haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
 (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 and parasitic infections.  
 Note: The sequence data for this patent did not form part of the

CC Printed specification, but was obtained in electronic format directly  
 CC from WPI at ftp.wpi.int/pub/published\_pct\_sequences.  
 XX Sequence 1963 BP; 622 A; 365 C; 394 G; 581 T; 1 other.

Query Match 84.0%; Score 16.8; DB 24; Length 1963;  
 Best Local Similarity 90.0%; Pred. No. 1e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GGAAGAGCCCTGACAGCCA 20  
 Db 960 GCAAGGCCCTGACAGCCA 979

RESULT 9  
 AAS40991  
 ID AAS40991 standard; cDNA; 2390 BP.  
 XX  
 AC AAS40991;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE cDNA encoding novel human enzyme polypeptide #207.

KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;  
 KW ligase; hyperproliferative disorder; immunodeficiency disorder;  
 KW autoimmune disorder; neurological disorder; metabolic disorder;  
 KW inflammatory disorder; cardiovascular disorder; reproductive disorder;  
 KW blood-related disorder; infectious disorder; gene therapy; cytostatic;  
 KW anti arthritic; neurotropic; anticoagulant; ss.

OS Homo sapiens.

XX WO200155301-A2.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01239.

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0205515.

PR 19-MAY-2000; 2000US-0209467.

PR 07-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216447.

PR 11-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 14-JUL-2000; 2000US-0217496.

PR 26-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 14-AUG-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225477.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226686.

PR 22-AUG-2000; 2000US-0227182.



RESULT 10  
AAFA4731  
ID AAF44731 standard; cDNA; 2463 BP.  
XX AC AAF44731;  
XX DT 27-MAR-2001 (first entry)  
XX DE Novel protein kinase cDNA, SEQ ID NO: 112.  
XX KW Human; mouse; protein kinase; antiarthritic; antiosteoporotic; osteoporathic;  
KW immunosuppressive; cardiant; renal; antiinflammatory; antiaethmatic;  
KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;  
KW cancer; autoimmune; cardiovascular disease; neurodegenerative disease;  
KW inflammatory pelvic disease; stroke; inflammatory bowel disease;  
XX OS Homo sapiens.  
XX PN WO200073469-A2.  
XX PD 07-DEC-2000.  
XX PF 26-MAY-2000; 2000WO-US14842.  
XX PR 28-MAY-1999; 99US-0136503.  
XX PA (SUGB-) SUGEN INC.  
XX P1 Plozman GD, Martinez R, Whyte D, Sudersanam S;  
XX WPI; 2001-032161/04.  
XX DR P-PSDB; AAB65702.  
XX PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
XX PT treating immune-related diseases and disorders, cardiovascular disease,  
XX BS neurodegenerative diseases and/or cancers -  
XX Example 1; Fig 2; 310p; English.  
XX CC The present sequence encodes a novel protein kinase. The nucleic acids  
CC and the protein kinases they encode may be used in the treatment and  
CC diagnosis of diseases associated with inappropriate kinase expression  
CC such as immune-related diseases and disorders, cardiovascular disease,  
CC neurodegenerative diseases and/or cancers. The nucleic acids and  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays. The kinase polypeptides may be used as antigens in the production  
CC of antibodies of kinase expression and activity. Anti-kinase antibodies  
CC and kinase antagonists may also be used to down regulate kinase  
CC activity. Diseases related to kinase expression and  
CC disorders, complications of organ transplantation, myocardial infarction,  
CC immune disorders, cardiomyopathies, strokes, renal failure,  
CC oxidative-stress related disorders, chronic inflammatory bowel disease,  
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,  
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and  
CC reproductive disorders.  
XX XX Sequence 2463 BP; 622 A; 686 C; 706 G; 447 T; 2 other;  
SQ

Query Match 84.0%; Score 16.8; DB 22; Length 2463;  
Best Local Similarity 90.0%; Pred. No. 1.1e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GGAAAGCCCTGACAGCA 20  
DB 2031 GGCAAGCCCTCACAAGCA 2050

ID AAS40990 standard; cDNA; 3085 BP.  
XX AC AAS40990;  
XX DT 17-DEC-2001 (first entry)  
XX DE cDNA encoding novel human enzyme polypeptide #206.  
XX KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;  
KW ligase; hyperproliferative disorder; immunodeficiency disorder;  
KW autoimmune disorder; neurological disorder; metabolic disorder;  
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;  
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;  
KW anti arthritic; neurotropic; anticoagulant; ss.  
XX OS Homo sapiens.  
XX PN WO200155301-A2.  
XX PD 02-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US01239.  
XX PR 31-JAN-2000; 2000US-0179065.  
XX PR 04-FEB-2000; 2000US-0180628.  
XX PR 24-FEB-2000; 2000US-0184664.  
XX PR 02-MAR-2000; 2000US-0186350.  
XX PR 16-MAR-2000; 2000US-0189874.  
XX PR 17-MAR-2000; 2000US-0190076.  
XX PR 18-APR-2000; 2000US-0198123.  
XX PR 19-MAY-2000; 2000US-0205515.  
XX PR 07-JUN-2000; 2000US-0209467.  
XX PR 28-JUN-2000; 2000US-0214886.  
XX PR 30-JUN-2000; 2000US-0215135.  
XX PR 07-JUL-2000; 2000US-0216647.  
XX PR 07-JUL-2000; 2000US-0216680.  
XX PR 11-JUL-2000; 2000US-0217487.  
XX PR 11-JUL-2000; 2000US-0217496.  
XX PR 26-JUL-2000; 2000US-0220963.  
XX PR 26-JUL-2000; 2000US-0220964.  
XX PR 14-AUG-2000; 2000US-0224518.  
XX PR 14-AUG-2000; 2000US-0224519.  
XX PR 14-AUG-2000; 2000US-0225213.  
XX PR 14-AUG-2000; 2000US-0225214.  
XX PR 14-AUG-2000; 2000US-0225266.  
XX PR 14-AUG-2000; 2000US-0225267.  
XX PR 14-AUG-2000; 2000US-0225268.  
XX PR 14-AUG-2000; 2000US-0225270.  
XX PR 14-AUG-2000; 2000US-0225447.  
XX PR 14-AUG-2000; 2000US-0225757.  
XX PR 14-AUG-2000; 2000US-0225758.  
XX PR 14-AUG-2000; 2000US-0225759.  
XX PR 18-AUG-2000; 2000US-0226279.  
XX PR 22-AUG-2000; 2000US-0226681.  
XX PR 22-AUG-2000; 2000US-0226688.  
XX PR 22-AUG-2000; 2000US-0227182.  
XX PR 23-AUG-2000; 2000US-0227009.  
XX PR 30-AUG-2000; 2000US-0228924.  
XX PR 01-SEP-2000; 2000US-0228924.  
XX PR 01-SEP-2000; 2000US-0229343.  
XX PR 01-SEP-2000; 2000US-0229344.  
XX PR 01-SEP-2000; 2000US-0229345.  
XX PR 05-SEP-2000; 2000US-0229509.  
XX PR 05-SEP-2000; 2000US-0229513.  
XX PR 06-SEP-2000; 2000US-0230437.  
XX PR 06-SEP-2000; 2000US-0230438.  
XX PR 08-SEP-2000; 2000US-0231242.  
XX PR 08-SEP-2000; 2000US-0231243.  
XX PR 08-SEP-2000; 2000US-0231244.  
XX PR 08-SEP-2000; 2000US-0231413.  
XX PR 08-SEP-2000; 2000US-0231414.  
XX PR 08-SEP-2000; 2000US-0232080.

[illegible][illegible]

OS Lactococcus sp.

Key Location/Qualifiers  
 FT promoter 1..153  
 FT -35\_signal 88..93  
 FT -10\_signal 110..116  
 FT RBS 136..144  
 FT CDS 154..180  
 FT /tag= d  
 FT /note= "5' end of coding sequence"

WO9738587-A1.

23-OCT-1997.

15-APR-1997; 97WO-NL00192.

15-APR-1996; 96EP-0200948.

(FRIE-) FRIESLAND BRADS BV.

Kok J, Mierau I, Van Den Burg L, Venema G;

WPI; 1997-535460/49.

Bio-transformation of a substrate using modified lactic acid bacterium - which contain a gene encoding altered neutral protease activity

Example 1; Fig 5; 61pp; English.

This sequence comprises lactococcal promoter P23, which was obtained from plasmid pGK223 by PCR amplification, and which has been inserted into PCR10 upstream of the Bacillus subtilis neutral protease npr gene (see AAT91284) to yield PCR12. Lactococcal CC used to control expression of the B. subtilis npr gene in vector constructs used for transformation of the B. subtilis npr gene in vector strains. The npr gene can be engineered to code for an enzyme (see AAT27278) that has altered activity, stability or specificity. CC neutral protease properties of the bacterium transformant allows the intended application. The method is especially used to tailor to an making.

Sequence 180 BP; 85 A; 22 C; 31 G; 42 T; 0 other;

Query Match 82.0%; Score 16.4; DB 18; Length 180;

Best Local Similarity 94.4%; Pred. No. 1.2e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 GAAAGCCCTGACAGCC 19

3 GAAAGCCCTGACAGCC 20

ABN69336

ABN69336 standard; DNA; 936 BP.

01-JUL-2002 (first entry)

Streptococcus polymnucleotide SEQ ID NO 6585.

Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

group A streptococcus; Streptococcus pyogenes; antibacterial; gene; antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.

OS Streptococcus pyogenes.

WO200234771-A2.

02-MAY-2002.

29-OCT-2001; 2001WO-GB04789.

27-OCT-2000; 2000GB-0026333.

24-NOV-2000; 2000GB-0028727.

07-MAR-2001; 2001GB-0005640.

(CHIR-) CHIRON SPA.

(GENO-) INST GENOMIC RES.

Telford J, Maignani V, Margarit Ros YI, Grandi G, Fraser C;

WPI; 2002-352536/38.

P-PSDB; ABP28705.

New Streptococcus protein for the treatment or prevention of infection or disease caused by Streptococcus bacteria, such as meningitis, and for detecting a compound that binds to the protein -

Claim 7; Page 3821; 4525pp; English.

The invention relates to a protein (ABP25413-ABP30895) from group B Streptococcus/GBS (Streptococcus agalactiae) or group A Streptococcus/GAS (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in the specification. The proteins have antibacterial and antiinflammatory activity. (i), nucleic acids encoding (i), ABN6044-ABN71526 and CC used in the treatment or prevention of infection or disease caused by Streptococcus bacteria, particularly S. agalactiae and S. pyogenes. CC biological sample. (i) is used to detect Streptococcus in a CC (i). A composition comprising (i) or a nucleic acid encoding (i) is used as a vaccine or diagnostic composition. The disease caused by Streptococcus that is prevented or treated may be meningitis. Nucleic acid encoding (i) may be used to recombinantly produce (i) and may be used in gene therapy. Antibodies to (i) are used for affinity CC chromatography, immunoassays, and distinguishing/identifying CC Streptococcus proteins.

Sequence 936 BP; 294 A; 180 C; 168 G; 294 T; 0 other;

Query Match 82.0%; Score 16.4; DB 24; Length 936;

Best Local Similarity 94.4%; Pred. No. 1.5e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 GAAAGCCCTGACAGCC 19

549 GAAAGCCCTGACAGAC 566

AAA26448

AAA26448 standard; CDNA; 2214 BP.

29-JUN-2000 (first entry)

Human secreted protein gene 20 SEQ ID NO:113.

Human; secreted protein; diagnosis; cytostatic; immunosuppressive;

antiHIV; antiinflammatory; nocitropic; neuroprotective; antiasthmatic; osteoporotic; antiarteritic; antibacterial; antidiabetic; antitubercular; antiparasitic; candidant; gene therapy; cancer; neurological disorder; immune disease; inflammation; blood disorder; tumour; ss.







Sun Aug 3 09:04:13 2003

US-09-834-291-6.rmp

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:05:17 ; Search time 5.18607 Seconds  
(Without alignments)  
795.924 Million cell updates/sec

Title: US-09-834-291-6  
Perfect score: 20  
Sequence: 1 ggaagagccctgacagcca 20  
Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues 28795534

Total number of hits satisfying chosen parameters:  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	US-09-834-291-6	Sequence 6, Appli
2	18.4	92.0	20	US-09-834-291-7	Sequence 7, Appli
3	18.4	92.0	20	US-09-834-291-10	Sequence 10, Appli
4	18.4	92.0	20	US-09-834-291-10	Sequence 24, Appli
5	18.4	92.0	20	US-09-834-291-32	Sequence 32, Appli
6	18.4	92.0	20	US-09-834-291-2	Sequence 2, Appli
7	18.4	92.0	20	US-09-834-291-3	Sequence 3, Appli
8	18.4	92.0	20	US-09-834-291-4	Sequence 4, Appli
9	18.4	92.0	20	US-09-834-291-1	Sequence 1, Appli
10	18.4	92.0	20	US-09-834-291-1	Sequence 250, App
11	16.4	84.0	13	US-10-027-632-134797	Sequence 1, Appli
12	16.4	82.0	14	US-10-156-761-1	Sequence 10989, A
13	15.8	82.0	14	US-09-918-995-10989	Sequence 15724, A
14	15.8	79.0	9	US-09-866-108-15724	Sequence 137408,
15	15.8	79.0	606	US-10-027-632-137408	
16	15.8	79.0	13	US-10-027-632-137408	

17	15.8	79.0	606	13	US-10-027-632-137409	Sequence 137409,
18	15.8	79.0	814	13	US-10-027-632-8488	Sequence 8488, Ap
19	15.8	79.0	907	13	US-10-027-632-137413	Sequence 137413,
20	15.8	79.0	907	13	US-10-027-632-137414	Sequence 137414,
21	15.8	79.0	4796	10	US-09-764-847-1627	Sequence 1627, Ap
22	15.8	79.0	4796	10	US-09-764-847-1628	Sequence 1628, Ap
23	15.8	79.0	10519	10	US-09-764-847-1628	Sequence 312, App
24	15.8	79.0	10519	10	US-09-764-847-1628	Sequence 1629, Ap
25	15.8	79.0	13337	14	US-09-764-847-1629	Sequence 1629, Ap
26	15.8	79.0	13337	14	US-10-091-483-112	Sequence 1629, Ap
27	15.8	79.0	13337	14	US-10-092-154-1629	Sequence 1734, Ap
28	15.8	79.0	21129	9	US-09-764-847-1734	Sequence 922, App
29	15.8	79.0	21129	11	US-09-764-847-1734	Sequence 922, App
30	15.8	79.0	31813	14	US-10-025-428-932	Sequence 366, App
31	15.8	79.0	31813	14	US-10-025-428-932	Sequence 366, App
32	15.4	77.0	508	10	US-09-833-263-366	Sequence 111067,
33	15.4	77.0	508	10	US-10-025-380-366	Sequence 111067,
34	15.4	77.0	654	13	US-10-027-632-111066	Sequence 111067,
35	15.4	77.0	654	13	US-10-027-632-111067	Sequence 111067,
36	15.4	77.0	654	13	US-10-027-632-111067	Sequence 111067,
37	15.4	77.0	789	14	US-10-198-846-4234	Sequence 4234, Ap
38	15.4	77.0	1119	14	US-10-001-857-68	Sequence 68, Appli
39	15.4	77.0	1253	13	US-10-231-426-3	Sequence 3, Appli
40	15.4	77.0	18540	14	US-10-105-637-1	Sequence 1, Appli
41	15.4	77.0	90442	10	US-09-933-267A-1	Sequence 8, Appli
42	15.4	77.0	465237	9	US-09-834-291-6	Sequence 9, Appli
43	15.2	76.0	20	9	US-09-834-291-9	
44	15.2	76.0	20	9	US-09-834-291-9	
45	15.2	76.0	20	9	US-09-834-291-9	

ALIGNMENTS

RESULT 1  
US-09-834-291-6 Application US/09834291  
Sequence 6, Appli  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIORITY APPLICATION NUMBER: PCT/DE99/03343  
PRIORITY FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-6  
Query Match 100.0%; Score 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0;  
DB 1 GGAAGAGCCCTGACAGCCA 20  
1 GGAAGAGCCCTGACAGCCA 20  
RESULT 2  
US-09-834-291-5 Application US/09834291  
Sequence 5, Appli  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter

APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: P53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 5  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-5

Query Match  
Best Local Similarity 92.0%; Score 18.4; DB 9; Length 20;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20  
DB 1 GGACAGCCCTGACAGCCA 20

RESULT 3  
US-09-834-291-7  
Sequence 7, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: P53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 7  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-7

Query Match  
Best Local Similarity 92.0%; Score 18.4; DB 9; Length 20;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20  
DB 1 GGAAGAGCCCTGACAGCCA 20

RESULT 4  
US-09-834-291-10  
Sequence 10, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: P53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343

PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 10  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-10

Query Match  
Best Local Similarity 92.0%; Score 18.4; DB 9; Length 20;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20  
DB 1 GGACAGCCCTGACAGCCA 20

RESULT 5  
US-09-834-291-24  
Sequence 24, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: P53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 24  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-24

Query Match  
Best Local Similarity 92.0%; Score 18.4; DB 9; Length 20;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20  
DB 1 GGACAGCCCTGACAGCCA 20

RESULT 6  
US-09-834-291-32  
Sequence 32, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: P53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 32  
LENGTH: 266

Sun Aug 3 09:04:13 2003

us-09-834-291-6.rnpb

TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-32

Query Match 92.0%; Score 18.4; DB 9; Length 266;  
Best Local Similarity 95.0%; Pred. No. 9.8;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20  
DB 160 GGACAGCCCTGACAGCCA 179

RESULT 7  
US-09-834-291-2  
Sequence 2, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Kramer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 2  
LENGTH: 720  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-2

QY 1 GGAAGAGCCCTGACAGCCA 20  
DB 160 GGACAGCCCTGACAGCCA 179

Query Match 92.0%; Score 18.4; DB 9; Length 720;  
Best Local Similarity 95.0%; Pred. No. 10;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20  
DB 160 GGACAGCCCTGACAGCCA 179

RESULT 8  
US-09-834-291-3  
Sequence 3, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Kramer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 3  
LENGTH: 2380  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-3

QY 1 GGAAGAGCCCTGACAGCCA 20  
DB 1820 GGACAGCCCTGACAGCCA 1839

RESULT 9  
US-09-834-291-4  
Sequence 4, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Kramer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2827  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-4

Query Match 92.0%; Score 18.4; DB 9; Length 2827;  
Best Local Similarity 95.0%; Pred. No. 10;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20  
DB 2267 GGACAGCCCTGACAGCCA 2286

RESULT 10  
US-09-834-291-1  
Sequence 1, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Kramer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 1  
LENGTH: 3212  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-1

Query Match 92.0%; Score 18.4; DB 9; Length 3212;  
Best Local Similarity 95.0%; Pred. No. 10;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGAGCCCTGACAGCCA 20  
DB 2659 GGACAGCCCTGACAGCCA 2678

RESULT 11

US-09-964-824A-250  
; Sequence 250, Application US/09964824A  
; Patent No. US20020102531A1  
; GENERAL INFORMATION:  
; APPLICANT: Horrikan, Stephen  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu  
; FILE REFERENCE: 689290-73  
; CURRENT APPLICATION NUMBER: US/09/964, 824A  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: US/60/236, 033  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US/60/236, 032  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US/60/236, 028  
; NUMBER OF SEQ ID NOS: 583  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 250  
; LENGTH: 1923  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-964-824A-250

Query Match  
Best Local Similarity 84.0%; Score 16.8; DB 10; Length 1923;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GGAAGCCCTGACAGCCA 20  
DB 1572 GGAAGCCCTGACAGCCA 1591

RESULT 12  
US-10-027-632-134797/c  
; Sequence 134797, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827,129  
; CURRENT APPLICATION NUMBER: US/10/027, 632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218, 006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198, 676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193, 483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185, 218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167, 363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156, 358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146, 002  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 134797  
; LENGTH: 824  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-134797

Query Match  
Best Local Similarity 82.0%; Score 16.4; DB 13; Length 824;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3 AAAAGCCCTGACAGCCA 20  
DB 694 AAAAGCCCTGACAGCCA 677

RESULT 13  
US-10-156-761-1/c  
; Sequence 1, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMTURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HAYTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156, 761  
; PRIOR FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 1  
; LENGTH: 9025608  
; TYPE: DNA  
; ORGANISM: Streptomyces avermitilis  
; FEATURE:  
; NAME/KEY: misc-feature  
; LOCATION: (4187715)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
US-10-156-761-1

Query Match  
Best Local Similarity 82.0%; Score 16.4; DB 14; Length 9025608;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GAAAGCCCTGACAGGCC 19  
DB 8367754 GAAAGCCCTGAAAGCC 8367757

RESULT 14  
US-09-918-995-10989  
; Sequence 10989, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918, 995  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235, 076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 10989  
; LENGTH: 438  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-918-995-10989

Query Match  
Best Local Similarity 79.0%; Score 15.8; DB 11; Length 438;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GAAAGCCCTGACAGCCA 20  
DB 383 GAAAGCCCTGACAGCCA 401

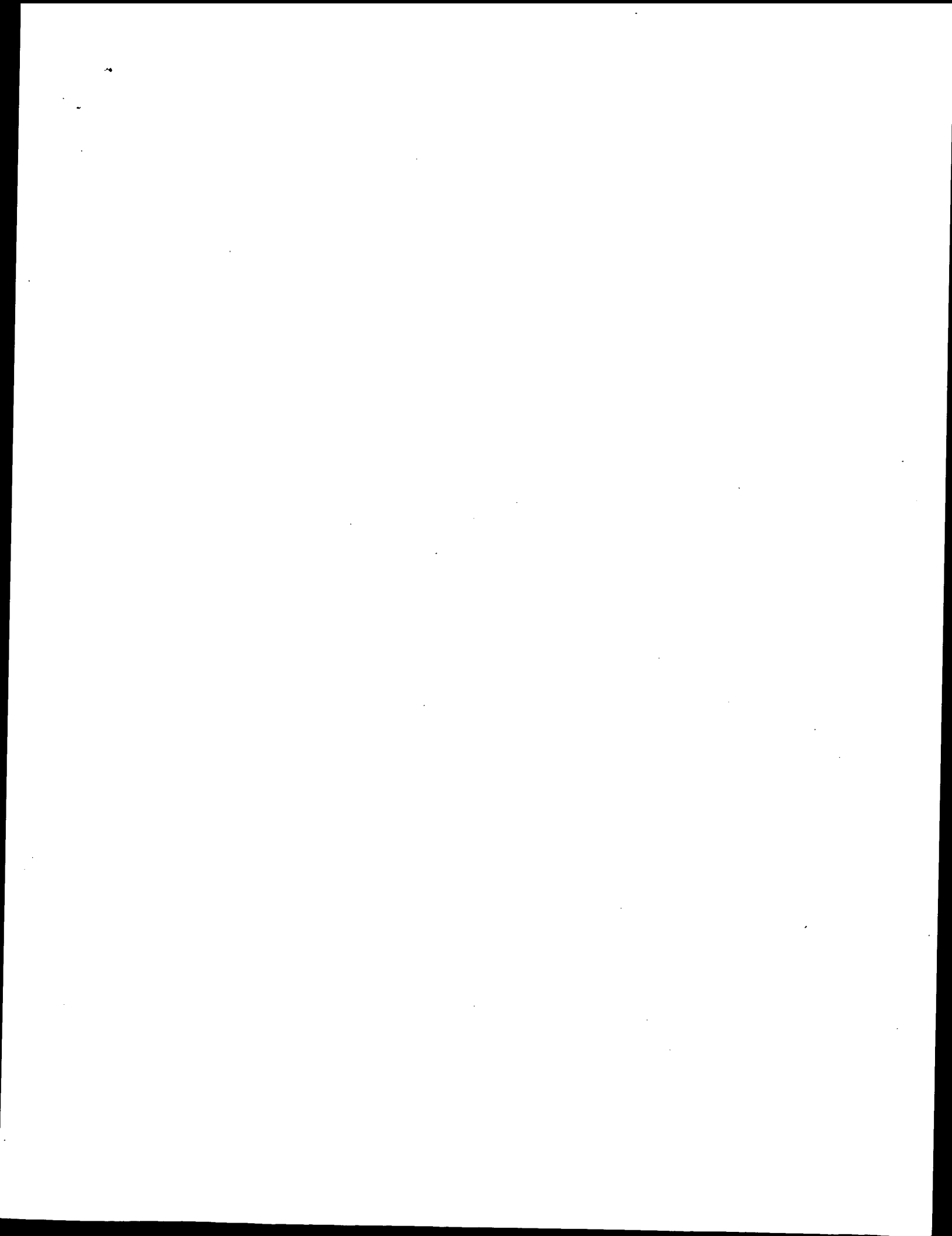
RESULT 15  
US-09-866-108-15724

```
; Sequence 15724, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharon G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 15724
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-866-108-15724

Query Match      79.0%; Score 15.8; DB 9; Length 500;
Best Local Similarity 89.5%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

QY 1 GGAAAGCCCTGACAGCC 19  
DB 383 GGGAAGCCCTGACAGGC 401

Search completed: August 1, 2003, 13:36:44  
Job time: 17.1861 secs



Sun Aug 3 09:04:12 2003

us-09-834-291-6.rml

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
August 1, 2003, 02:21:37 ; Search time 1.65066 Seconds  
(without alignments)  
5347.959 Million cell updates/sec

Run on: August 1, 2003, 02:21:37 ; Search time 1.65066 Seconds  
(without alignments)  
5347.959 Million cell updates/sec

Title: US-09-834-291-6  
Perfect score: 20  
Sequence: 1 ggaagccctgacagcca 20

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 563978 seqs, 220691566 residues  
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgnt2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgnt2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
3: /cgnt2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
4: /cgnt2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
5: /cgnt2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
6: /cgnt2\_6/ptodata/1/ina/5A\_COMB.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.8	84.0	1923	US-09-220-132-85	Sequence 85, Appl
2	15.8	79.0	72	US-08-489-847-113	Sequence 113, Appl
3	15.8	79.0	72	US-08-489-847-113	Sequence 93, Appl
4	15.8	79.0	72	US-08-489-847-113	Sequence 49, Appl
5	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1434, A
6	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1416, A
7	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1390, A
8	15.2	76.0	1107	US-09-252-991A-1434	Sequence 88, Appl
9	15.2	76.0	1107	US-09-252-991A-1434	Sequence 88, Appl
10	15.2	76.0	1107	US-09-252-991A-1434	Sequence 2, Appl
11	15.2	76.0	1107	US-09-252-991A-1434	Sequence 17, Appl
12	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1, Appl
13	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1, Appl
14	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1, Appl
15	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1, Appl
16	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1, Appl
17	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1, Appl
18	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1, Appl
19	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1, Appl
20	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1, Appl
21	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1, Appl
22	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1, Appl
23	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1, Appl
24	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1, Appl
25	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1, Appl
26	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1, Appl
27	15.2	76.0	1107	US-09-252-991A-1434	Sequence 1, Appl

28	14.4	72.0	17138	US-09-220-048-3	Sequence 3, Appl
29	14.4	72.0	36181	US-08-311-711A-120	Sequence 120, Appl
30	14.4	71.0	90	US-08-155-171B-3	Sequence 3, Appl
31	14.2	71.0	90	US-08-435-998-3	Sequence 86, Appl
32	14.2	71.0	218	US-09-188-930-86	Sequence 2050, Appl
33	14.2	71.0	218	US-09-312-283C-86	Sequence 25, Appl
34	14.2	71.0	315	US-08-107-532A-2050	Sequence 25, Appl
35	14.2	71.0	384	US-08-799-464A-25	Sequence 18, Appl
36	14.2	71.0	387	US-08-855-526B-15	Sequence 15, Appl
37	14.2	71.0	387	US-08-855-526B-15	Sequence 25, Appl
38	14.2	71.0	387	US-08-855-526B-15	Sequence 25, Appl
39	14.2	71.0	387	US-08-855-526B-15	Sequence 25, Appl
40	14.2	71.0	387	US-08-855-526B-15	Sequence 25, Appl
41	14.2	71.0	387	US-08-855-526B-15	Sequence 25, Appl
42	14.2	71.0	387	US-08-855-526B-15	Sequence 25, Appl
43	14.2	71.0	387	US-08-855-526B-15	Sequence 25, Appl
44	14.2	71.0	387	US-08-855-526B-15	Sequence 25, Appl
45	14.2	71.0	387	US-08-855-526B-15	Sequence 25, Appl

ALIGNMENTS

RESULT 1  
US-09-220-132-85  
Sequence 85, Application US/09220132  
Patent No. 6506607  
GENERAL INFORMATION: Andrew W.  
APPLICANT: Shyjan, METHODS AND COMPOSITIONS FOR THE IDENTIFICATION OF PROSTATE CANCER  
TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 07334-074001  
CURRENT FILING DATE: 1998-12-23  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: US 60/079,303  
PRIOR FILING DATE: 1997-12-24  
PRIOR APPLICATION NUMBER: US 60/068,821  
NUMBER OF SEQ ID NOS: 191  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 85  
LENGTH: 1923  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-220-132-85

Query Match 84.0%; Score 16.8; DB 4; Length 1923;  
Best Local Similarity 90.0%; Pred. No. 12;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGCCCTGACAGCCA 20  
DB 1572 GGAAGCCCTGACAGCCA 1591  
RESULT 2  
US-09-489-847-113  
Sequence 113, Application US/09489847  
Patent No. 6476195  
GENERAL INFORMATION: Rosen et al  
APPLICANT: Rosen et al  
TITLE OF INVENTION: 98 Human Secreted Proteins  
FILE REFERENCE: P2031P1  
CURRENT FILING DATE: 2000-01-24  
PRIOR FILING DATE: 1999-07-29  
PRIOR APPLICATION NUMBER: PCT/US99/17130  
PRIOR FILING DATE: 1998-07-30  
PRIOR APPLICATION NUMBER: 60/094,657  
PRIOR FILING DATE: 1998-08-05  
PRIOR APPLICATION NUMBER: 60/096,319  
PRIOR FILING DATE: 1998-08-12

EARLIER APPLICATION NUMBER: 60/095,454  
EARLIER FILING DATE: 1998-08-06  
EARLIER APPLICATION NUMBER: 60/095,455  
NUMBER OF SEQ ID NOS: 376  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 113  
LENGTH: 2214  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (289)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-489-847-113

Query Match  
Best Local Similarity 80.0%; Score 16; DB 4; Length 2214;  
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 GAAAGCCCTGACAGCA 20  
Db 305 GGGARAGCCTGACAGCA 324

RESULT 3  
US-09-276-531-93  
Sequence 93, Application US/09276531  
Patent No. 6183968  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Yue, Henry  
APPLICANT: Reddy, Roopa  
APPLICANT: Guegler, Karl J.  
APPLICANT: Baughn, Mariah R.  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING  
NUMBER OF SEQUENCES: 134  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/276,531  
FILING DATE: Herewith  
CLASSIFICATION:  
Prior Application:  
APPLICATION DATA:  
APPLICATION NUMBER: 60/079,677  
FILING DATE: March 27, 1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Lynn E. Murty, Ph.D.  
REGISTRATION NUMBER: 42,918  
TELECOMMUNICATION INFORMATION: PA-0008 US  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 93:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 72 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

IMMEDIATE SOURCE:  
LIBRARY: BRSTU08  
CLONE: 1729463  
US-09-276-531-93

Query Match  
Best Local Similarity 79.0%; Score 15.8; DB 3; Length 72;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 2 GAAAGCCCTGACAGCA 20  
Db 32 GAAAGCCCTGACAGCA 50

RESULT 4  
US-08-894-997-49  
Sequence 49, Application US/08894997A  
Patent No. 6270990  
GENERAL INFORMATION:  
APPLICANT: Anderson, David J  
APPLICANT: Schoenheit, Christopher J  
TITLE OF INVENTION: NEURON-RESTRICTIVE SILENCER FACTOR  
FILE REFERENCE: 17810-502 NRSP  
CURRENT APPLICATION NUMBER: US/08/894,997A  
EARLIER FILING DATE: 1998-01-06  
EARLIER APPLICATION NUMBER: PCT/US96/02817  
EARLIER FILING DATE: 1996-03-01  
NUMBER OF SEQ ID NOS: 55  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 49  
LENGTH: 4057  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: gene  
LOCATION: (1)..(4057)  
OTHER INFORMATION: Human NSRP  
US-08-894-997-49

Query Match  
Best Local Similarity 79.0%; Score 15.8; DB 3; Length 4057;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 2 GAAAGCCCTGACAGCA 20  
Db 1895 GAAAGCCCTGTAAGCA 1913

RESULT 5  
US-09-252-991A-14234  
Sequence 14234, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
PRIOR FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
NUMBER OF SEQ ID NOS: 1998-07-27  
SEQ ID NO 14234  
LENGTH: 738  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-14234

Query Match  
76.0%; Score 15.2; DB 4; Length 738;



Best Local Similarity 85.0%; Pred. No. 67;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGAAAGCCTGACAGCCA 20  
160 GGAAAGCCTGACAGCCA 179

RESULT 6  
US-09-252-991A-14146  
Sequence 14146, Application US/09252991A

PATENT No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT FILING DATE: 1999-02-18  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-07-27  
PRIOR APPLICATION NUMBER: US 60/094,190  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 14146  
LENGTH: 1107  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-14146

Query Match 76.0%; Score 15.2; DB 4; Length 1107;  
Best Local Similarity 85.0%; Pred. No. 72;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGAAAGCCTGACAGCCA 20  
597 GGAAAGCCTGACAGCCA 616

RESULT 7  
US-09-252-991A-13990/c  
Sequence 13990, Application US/09252991A

PATENT No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT FILING DATE: 1999-02-18  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-07-27  
PRIOR APPLICATION NUMBER: US 60/094,190  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 13990  
LENGTH: 1521  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-13990

Query Match 76.0%; Score 15.2; DB 4; Length 1521;  
Best Local Similarity 85.0%; Pred. No. 77;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGAAAGCCTGACAGCCA 20  
739 GGAAAGCCTGACAGCCA 720

RESULT 8  
US-09-232-200-88/c  
Sequence 88, Application US/09232200A

PATENT No. 6288213  
GENERAL INFORMATION:

APPLICANT: Stahl, Andreas  
APPLICANT: Hirsch, David J.  
APPLICANT: Lodish, Harvey F.  
APPLICANT: Glimo, Ruth E.  
APPLICANT: Tartaglia, Louis A.  
TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
FILE REFERENCE: WH197-2193MB  
CURRENT FILING DATE: 1999-01-14  
PRIOR FILING DATE: 1998-01-15  
PRIOR APPLICATION NUMBER: 60/093,491  
PRIOR FILING DATE: 1998-07-20  
PRIOR APPLICATION NUMBER: 60/110,941  
EARLIER FILING DATE: 1998-12-04  
NUMBER OF SEQ ID NOS: 105  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 88  
LENGTH: 1872  
TYPE: DNA  
ORGANISM: Saccharomyces cerevisiae  
US-09-232-200-88

Query Match 76.0%; Score 15.2; DB 3; Length 1872;  
Best Local Similarity 85.0%; Pred. No. 80;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGAAAGCCTGACAGCCA 20  
46 GGAAATCTTGACAGCCA 27

RESULT 9  
US-09-232-197-88/c  
Sequence 88, Application US/09232197A

PATENT No. 6300096  
GENERAL INFORMATION:  
APPLICANT: Stahl, Andreas  
APPLICANT: Hirsch, David J.  
APPLICANT: Lodish, Harvey F.  
APPLICANT: Glimo, Ruth E.  
APPLICANT: Tartaglia, Louis A.  
TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
FILE REFERENCE: WH197-2193MA  
CURRENT FILING DATE: 1999-01-14  
PRIOR FILING DATE: 1998-01-15  
PRIOR APPLICATION NUMBER: 60/071,374  
PRIOR FILING DATE: 1998-07-20  
PRIOR APPLICATION NUMBER: 60/093,491  
PRIOR FILING DATE: 1998-12-04  
NUMBER OF SEQ ID NOS: 105  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 88  
LENGTH: 1872  
TYPE: DNA  
ORGANISM: Saccharomyces cerevisiae  
US-09-232-197-88

Query Match 76.0%; Score 15.2; DB 4; Length 1872;  
Best Local Similarity 85.0%; Pred. No. 80;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGAAAGCCTGACAGCCA 20  
46 GGAAATCTTGACAGCCA 27

RESULT 10  
US-09-232-201-88/c

```

; Sequence 88, Application US/09232201A
; Patent No. 6348321
; GENERAL INFORMATION:
; APPLICANT: Stahl, Andreas
; APPLICANT: Hirsch, David J.
; APPLICANT: Lodish, Harvey F.
; APPLICANT: Gimeno, Ruth E.
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
; FILE REFERENCE: WHI97-21D3MC
; CURRENT APPLICATION NUMBER: US/09/232, 201A
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/071,374
; EARLIER FILING DATE: 1998-01-15
; EARLIER APPLICATION NUMBER: 60/093,491
; EARLIER FILING DATE: 1998-07-20
; EARLIER APPLICATION NUMBER: 60/110,941
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 88
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-232-201-88

```

```

Query Match
Best Local Similarity 76.0%; Score 15.2; DB 4; Length 1872;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 GGAAGAGCCCTGACAGCA 20
DB 46 GGAAGATCTTGACAGCA 27

```

```

RESULT 11
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "a" bases at various positions throughout the sequence
US-09-103-840A-2

```

```

Query Match
Best Local Similarity 76.0%; Score 15.2; DB 3; Length 4403765;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 GGAAGAGCCCTGACAGCA 20
DB 4172848 GGAAGAGCCCTGATATATCA 4172867

```

```

RESULT 12
US-09-103-840A-1
; Sequence 1, Application US/09103840A

```

```

; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

```

```

Query Match
Best Local Similarity 76.0%; Score 15.2; DB 3; Length 4411529;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 GGAAGAGCCCTGACAGCA 20
DB 4180600 GGAAGAGCCCTGATATATCA 4180619

```

```

RESULT 13
US-09-634-238-137
; Sequence 137, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
; APPLICANT: Giem, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
; APPLICANT: Christensen, Anna C.
; APPLICANT: Holland, Rose
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; FILE REFERENCE: 11000.1043U1
; CURRENT APPLICATION NUMBER: US/09/634,238
; CURRENT FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 137
; LENGTH: 2221
; TYPE: DNA
; ORGANISM: Lactobacillus rhamnosus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ... (2221)
; OTHER INFORMATION: n = A,T,C or G
US-09-634-238-137

```

```

Query Match
Best Local Similarity 74.0%; Score 14.8; DB 4; Length 2221;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 GGAAGAGCCCTGACAGC 18
DB 283 GGAAGAGCCCTGACAGC 300

```

```

RESULT 14
US-09-135-782-1/C
; Sequence 1, Application US/09135782

```

Patent No. 6027929  
GENERAL INFORMATION:  
APPLICANT: Xu, Shuang-yong  
TITLE OF INVENTION: Method For Cloning And Producing The NspI Restriction  
TITLE OF INVENTION: Endonuclease In E. coli And Purification Of The  
TITLE OF INVENTION: Recombinant NspI Restriction Endonuclease  
FILE REFERENCE: NEB-143  
CURRENT APPLICATION NUMBER: US/09/135,782  
CURRENT FILING DATE: 1998-08-18  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO: 1  
LENGTH: 1194  
TYPE: DNA  
ORGANISM: No. 6027929LOC SP.  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1191)  
US-09-135-782-1

Query Match 72.0%; Score 14.4; DB 3; Length 1194;  
Best Local Similarity 93.8%; Pred. No. 1.9e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GAAAGCCCTGACAG 17  
DB 422 GAAAGCCCTGACAG 407

RESULT 15  
US-08-240-372-1  
Sequence 1, Application US/08240372  
Patent No. 5741665  
GENERAL INFORMATION:  
APPLICANT: KATO, ELIE K.  
APPLICANT: STUART, W. DORSEY  
TITLE OF INVENTION: LIGHT-REGULATED PROMOTERS FOR PRODUCTION  
TITLE OF INVENTION: OF HETEROLOGOUS PROTEINS IN FILAMENTOUS FUNGI  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W.  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/240,372  
FILING DATE: 10-MAY-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 3918-0003.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1237 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-240-372-1

Query Match 72.0%; Score 14.4; DB 1; Length 1237;  
Best Local Similarity 93.8%; Pred. No. 1.9e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 AAGCCCTGACAGCCA 20  
DB 365 AAGCCCTGACAGACA 380

Search completed: August 1, 2003, 08:37:08  
Job time: 13.6507 secs



Sun Aug 3 09:03:40 2003

US-09-834-291-10.rge

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:37:27 ; Search time 78.1573 Seconds  
(without alignments)  
10468.541 Million cell updates/sec

Title: US-09-834-291-10

Perfect score: 1 ggaacagccctgacacgca 20

Sequence: IDENTITY\_NUC

Scoring table: Gapop 10.0, Gapext 1.0

Searched: 2688711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenBank:\*

1: gb\_ba:\*

2: gb\_hcg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_ov:\*

22: em\_pat:\*

23: em\_ph:\*

24: em\_pl:\*

25: em\_ro:\*

26: em\_sts:\*

27: em\_sy:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_hcg\_hum:\*

31: em\_hcg\_inv:\*

32: em\_hcg\_other:\*

33: em\_hcg\_mus:\*

34: em\_hcg\_pln:\*

35: em\_hcg\_rod:\*

36: em\_hcg\_vir:\*

37: em\_hcg:\*

38: em\_by:\*

39: em\_hcgo\_hum:\*

40: em\_hcgo\_mus:\*

41: em\_hcgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match	Query Length	DB ID	Description
1	20	100.0	20	6	AX026093
2	20	100.0	20	6	AX026098
3	20	100.0	20	6	AX026112
4	20	100.0	266	6	AX026120
5	20	100.0	266	6	HS4011034
6	20	100.0	720	6	AX026090
7	20	100.0	2380	6	AX026091
8	20	100.0	2827	6	AX026092
9	20	100.0	3212	6	AX026089
10	20	100.0	45121	6	AX695635
11	20	100.0	187313	9	AL157394
12	18.4	92.0	167412	2	AL356352
13	18.4	92.0	192657	10	AL672160
14	18.4	92.0	193352	2	EX294664
15	18.4	92.0	199776	9	AC011499
16	18.4	92.0	204012	10	AL732613
17	18.4	90.0	81200	2	AC022086
18	18	90.0	145667	2	AL662817
19	18	90.0	214440	2	AC091422
20	18	90.0	215973	6	BF444228
21	18	87.0	80961	2	AL399977
22	17.4	87.0	80961	2	AL399959
23	17.4	87.0	163933	5	AL589763
24	17.4	87.0	166772	9	AL587943
25	17.4	87.0	175770	2	AC127943
26	17.4	87.0	185049	9	AL365495
27	17.4	87.0	186325	2	AC074244
28	17.4	87.0	189485	2	AC141268
29	17.4	87.0	193347	2	AC125884
30	17.4	87.0	203246	5	AL954133
31	17.4	87.0	206373	2	AL954133
32	17.4	87.0	223831	2	AC115292
33	17.4	87.0	228659	2	AC116393
34	17.4	87.0	230097	2	AC126053
35	17.4	87.0	238103	2	AC099140
36	17.4	87.0	258131	2	AC135751
37	17.4	87.0	289818	2	AC108332
38	17.4	85.0	236723	2	AC112477
39	17	85.0	240214	2	AC126871
40	17	85.0	257842	2	AC105569
41	17	85.0	262681	2	AC128975
42	17	85.0	280469	2	AC121541
43	16.8	84.0	20	6	AX026095
44	16.8	84.0	20	6	AX026097

ALIGNMENTS

RESULT 1

AX026093

LOCUS AX026093

DEFINITION Sequence 5 from Patent DE19847779.

ACCESSION AX026093

VERSION AX026093.1 GI:10187524

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Mueller-Schilling, M., Kramer, P. and Oren, M.

Novel receptor dna useful for identifying apoptosis-modulating

substances potentially useful for cancer chemotherapy

20 bp DNA linear PAT 16-SEP-2000

## JOURNAL

Patent: DE 19847779-C 5 03-FEB-2000;  
DEUTSCHES KREBSFORSCH (DE)

## FEATURES

Location/Qualifiers  
1..20

BASE COUNT  
ORIGIN 7 a 7 c 5 g 1 t

Query Match  
Best Local Similarity 100.0%; Score 20; DB 6; Length 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GGACAGCCCTGACAGCCA 20  
1 GGACAGCCCTGACAGCCA 20

## RESULT 2

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## FEATURES

## source

BASE COUNT  
ORIGIN 7 a 7 c 5 g 1 t

Query Match  
Best Local Similarity 100.0%; Score 20; DB 6; Length 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GGACAGCCCTGACAGCCA 20  
1 GGACAGCCCTGACAGCCA 20

## RESULT 3

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## FEATURES

## source

1..20  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT  
ORIGIN 7 a 7 c 5 g 1 t

Query Match  
Best Local Similarity 100.0%; Score 20; DB 6; Length 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GGACAGCCCTGACAGCCA 20  
1 GGACAGCCCTGACAGCCA 20

## RESULT 4

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## FEATURES

## source

BASE COUNT  
ORIGIN 49 a 72 c 110 g 35 t

Query Match  
Best Local Similarity 100.0%; Score 20; DB 6; Length 266;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GGACAGCCCTGACAGCCA 20  
160 GGACAGCCCTGACAGCCA 179

## RESULT 5

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## FEATURES

## source

1..20  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

1..20  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

FEATURES GERMANY Location/Qualifiers  
 source 1..266 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 gene 159..178 /gene="CD95"  
 159..178 /gene="CD95"  
 enhancer /evidence=experimental  
 BASE COUNT 49 a 72 c 110 g 35 t  
 ORIGIN  
 Query Match 100.0%; Score 20; DB 9; Length 266;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGACAGCCCTGACAGCCA 20  
 160 GGACAGCCCTGACAGCCA 179  
 Db  
 RESULT 6 AX026090 720 bp DNA linear PAT 16-SEP-2000  
 LOCUS AX026090 Sequence 2 from Patent DE19847779.  
 DEFINITION AX026090  
 ACCESSION AX026090.1 GI:10187521  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 Mueller-Schilling, M., Krammer, P. and Oren, M.  
 Novel receptor dna useful for identifying apoptosis-modulating  
 substances potentially useful for cancer chemotherapy  
 Patent: DE 19847779-C 2 03-FEB-2000;  
 JOURNAL DEUTSCHES KREBSFORSCH (DE)  
 FEATURES Location/Qualifiers  
 source 1..720 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 142 a 181 c 216 g 181 t  
 ORIGIN  
 Query Match 100.0%; Score 20; DB 6; Length 720;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGACAGCCCTGACAGCCA 20  
 160 GGACAGCCCTGACAGCCA 179  
 Db  
 RESULT 7 AX026091 2380 bp DNA linear PAT 16-SEP-2000  
 LOCUS AX026091 Sequence 3 from Patent DE19847779.  
 DEFINITION AX026091  
 ACCESSION AX026091.1 GI:10187522  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 Mueller-Schilling, M., Krammer, P. and Oren, M.  
 Novel receptor dna useful for identifying apoptosis-modulating  
 substances potentially useful for cancer chemotherapy  
 Patent: DE 19847779-C 3 03-FEB-2000;  
 JOURNAL

FEATURES DEUTSCHES KREBSFORSCH (DE) Location/Qualifiers  
 source 1..2380 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 579 a 595 c 568 g 638 t  
 ORIGIN  
 Query Match 100.0%; Score 20; DB 6; Length 2380;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGACAGCCCTGACAGCCA 20  
 1820 GGACAGCCCTGACAGCCA 1839  
 Db  
 RESULT 8 AX026092 2827 bp DNA linear PAT 16-SEP-2000  
 LOCUS AX026092 Sequence 4 from Patent DE19847779.  
 DEFINITION AX026092  
 ACCESSION AX026092.1 GI:10187523  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 Mueller-Schilling, M., Krammer, P. and Oren, M.  
 Novel receptor dna useful for identifying apoptosis-modulating  
 substances potentially useful for cancer chemotherapy  
 Patent: DE 19847779-C 4 03-FEB-2000;  
 JOURNAL DEUTSCHES KREBSFORSCH (DE)  
 FEATURES Location/Qualifiers  
 source 1..2827 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 728 a 676 c 657 g 766 t  
 ORIGIN  
 Query Match 100.0%; Score 20; DB 6; Length 2827;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGACAGCCCTGACAGCCA 20  
 2267 GGACAGCCCTGACAGCCA 2286  
 Db  
 RESULT 9 AX026089 3212 bp DNA linear PAT 16-SEP-2000  
 LOCUS AX026089 Sequence 1 from Patent DE19847779.  
 DEFINITION AX026089  
 ACCESSION AX026089.1 GI:10187520  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 Mueller-Schilling, M., Krammer, P. and Oren, M.  
 Novel receptor dna useful for identifying apoptosis-modulating  
 substances potentially useful for cancer chemotherapy  
 Patent: DE 19847779-C 1 03-FEB-2000;  
 JOURNAL DEUTSCHES KREBSFORSCH (DE)  
 FEATURES Location/Qualifiers  
 source 1..3212 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"

BASE COUNT 778 a -784 c 809 g 841 t  
 ORIGIN

Query Match  
 Best Local Similarity 100.0%; Score 20; DB 6; Length 3212;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGACAGCCCTGACAGCCA 20  
 DB 2659 GGACAGCCCTGACAGCCA 2678

RESULT 10  
 LOCUS AX695635 45121 bp DNA linear PAT 31-MAR-2003  
 DEFINITION Sequence 1262 from Patent WO03008583.  
 ACCESSION AX695635  
 VERSION AX695635.1 GI:29418787  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Morris, D.W. and Engelhard, E.K.  
 TITLE Novel compositions and methods for cancer  
 JOURNAL Patent: WO 03008583-A 1262 30-JAN-2003;  
 FEATURES Sagres Discovery (US)  
 SOURCE Location/Qualifiers  
 1. .45121  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 13226 a 8836 c 9010 g 14049 t  
 ORIGIN

Query Match  
 Best Local Similarity 100.0%; Score 20; DB 6; Length 45121;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGACAGCCCTGACAGCCA 20  
 DB 10653 GGACAGCCCTGACAGCCA 10672

RESULT 11  
 LOCUS AL157394 187313 bp DNA linear PRI 22-AUG-2001  
 DEFINITION Human DNA sequence from clone Rpl1-399019 on chromosome 10,  
 complete sequence.  
 ACCESSION AL157394  
 VERSION AL157394.15 GI:15384622  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS 1 (bases 1 to 187313)  
 TITLE Direct Submission  
 JOURNAL Submitted (22-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 requests: clonerequest@sanger.ac.uk  
 CE10 15A, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 On Aug 31, 2001 this sequence version replaced gi:14161146.  
 Where differences are found these are annotated as variations.  
 Together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 \*regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SwissProt; Tr: TrEMBL; Wp: WormPEP; information on the WormPEP database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep)  
 This sequence is the entire insert of clone Rpl1-399019 The true left end of clone Rpl1-496H23 is at 166408 in this sequence. The true right end of clone Rpl1-30415 is at 18704 in this sequence.

FEATURES  
 Source  
 1. .187313  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="10"  
 /clone="Rpl1-399019"  
 /clone\_11b="Rpl1-11.2"  
 /note="Sequence confirmed by AC015461 sequenced by W1BR."  
 100119. 100156  
 /note="Sequence confirmed by AC015461 sequenced by W1BR."  
 100157. 100198  
 /note="Sequence confirmed by AC015461 sequenced by W1BR."  
 105808. 105972  
 /note="Sequence confirmed by AC015461 sequenced by W1BR."  
 105973. 105989  
 /note="Sequence confirmed by AC015461 sequenced by W1BR."  
 BASE COUNT 55669 a 36398 c 36888 g 58358 t  
 ORIGIN

Query Match  
 Best Local Similarity 100.0%; Score 20; DB 9; Length 187313;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGACAGCCCTGACAGCCA 20  
 DB 144628 GGACAGCCCTGACAGCCA 144647

RESULT 12  
 LOCUS AX026094 20 bp DNA linear PAT 16-SEP-2000  
 DEFINITION Sequence 6 from Patent DE19847779.  
 ACCESSION AX026094  
 VERSION AX026094.1 GI:10187525  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.  
 TITLE Novel receptor dna useful for identifying apoptosis-modulating  
 JOURNAL Patent: DE 19847779-C 6 03-FEB-2000;  
 DEUTSCHES KREBSFORSCH (DE)  
 FEATURES location/Qualifiers  
 1. .20  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 8 a 6 c 5 g 1 t  
 ORIGIN



Query Match 92.0%; Score 18.4; DB 6; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 4.5e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGACACGCTGACACCA 20  
 1 CGAAAGCCTGACACCA 20

RESULT 13  
 AL556352/c  
 LOCUS 167412 bp DNA 1linear HTG 21-OCT-2001  
 DEFINITION Homo sapiens chromosome 1 clone RP5-859H16, \*\*\* SEQUENCING IN  
 PROGRESS \*\*\*, 20 unordered pieces.  
 ACCESSION AL556352 GI:16304915  
 VERSION AL556352.8  
 KEYWORDS HTG: HTGS PHASE1; HTGS-CANCELLED.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
 AUTHORS McLay, K.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 requests: clonerequests@sanger.ac.uk  
 On Oct 21, 2001 this sequence version replaced gi:12539659.  
 COMMENT Genom Center

Center: Sanger Centre  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: humquerry@sanger.ac.uk  
 Project Information  
 Center project name: dJ859H16  
 Summary Statistics  
 Assembly program: XGAP4; version 4.5  
 Sequencing vector: Plasmid; L08752; 100% of reads Chemistry:  
 Chemistry: Dye-terminator ET-amersham; 26% of reads Chemistry:  
 Dye-terminator Big Dye; 73% of reads  
 Consensus quality: 15949 bases at least Q40  
 Consensus quality: 162148 bases at least Q30  
 Consensus quality: 163926 bases at least Q20  
 Consensus quality: sum-of-contigs  
 Insert size: 165512; sum-of-contigs  
 Insert size: 206683; 10.2% error; agarose-fp  
 Quality coverage: 4.08x in Q20 bases; sum-of-contigs Quality  
 coverage: 3.39x in Q20 bases; agarose-fp

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 20 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of 'N', but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 12019: contig of 12019 bp in length  
 \* 12020 12119: gap of 100 bp  
 \* 12120 32657: contig of 20536 bp in length  
 \* 32658 32757: gap of 100 bp  
 \* 32758 36105: contig of 3348 bp in length  
 \* 36106 36205: gap of 100 bp  
 \* 36206 41069: contig of 4864 bp in length  
 \* 41070 41169: gap of 100 bp  
 \* 41170 51754: contig of 10485 bp in length  
 \* 51755 62397: contig of 10643 bp in length  
 \* 62398 73595: gap of 100 bp  
 \* 73596 83277: contig of 11098 bp in length  
 \* 83278 83377: gap of 100 bp  
 \* 83378 83377: gap of 100 bp

## FEATURES

source 1. 167412  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /clone="RP5-859H16"  
 /clone\_1fb="RP5-859H16-5"  
 /clone\_2fb="RP5-859H16-5"  
 1. 12019  
 /note="assembly\_fragment:00380  
 clone\_end:866  
 vector\_side:left"  
 12120..32657  
 /note="assembly\_fragment:00114  
 fragment\_chain:1"  
 32758..36105  
 /note="assembly\_fragment:00063  
 fragment\_chain:1"  
 36206..41069  
 /note="assembly\_fragment:01365  
 fragment\_chain:1"  
 41170..51754  
 /note="assembly\_fragment:01782  
 fragment\_chain:1"  
 51755..62397  
 /note="assembly\_fragment:01003  
 fragment\_chain:1"  
 62498..73595  
 /note="assembly\_fragment:00248  
 fragment\_chain:2"  
 73696..83277  
 /note="assembly\_fragment:01125  
 fragment\_chain:2"  
 83378..98845  
 /note="assembly\_fragment:01780  
 fragment\_chain:2"  
 98946..101173  
 /note="assembly\_fragment:01201  
 fragment\_chain:2"  
 101274..103420  
 /note="assembly\_fragment:01704  
 fragment\_chain:3"  
 103521..109720  
 /note="assembly\_fragment:00350  
 fragment\_chain:3"  
 109821..114256  
 /note="assembly\_fragment:01435  
 fragment\_chain:3"  
 114357..119627  
 /note="assembly\_fragment:00291  
 fragment\_chain:3"

```

misc_feature      fragment chain:4"
119728..128350
/nc="assembly_fragment:01405
fragment_chain:4"
128451..140183
/nc="assembly_fragment:00901
fragment_chain:5"
140284..147703
/nc="assembly_fragment:00215
fragment_chain:5"
147804..158100
/nc="assembly_fragment:00322"
158201..161982
/nc="assembly_fragment:00595"
162083..167412
/nc="assembly_fragment:00872"
BASS COUNT      41496 a 41224 c 42341 g 40447 t 1904 others
ORIGIN
Query Match
Best Local Similarity 92.0%; Score 18.4; DB 2; Length 167412;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Db 143124 GTCACAGCCCTGACAGCA 143105
1
|||||
1 GGACAGCCCTGACAGCA 20

```

```

RESULT 14
AL672160      192657 bp      DNA      linear      ROD 17-APR-2003
LOCUS         Mouse DNA sequence from clone RP23-144K18 on chromosome 4, complete
DEFINITION
ACCESSION     AL672160
VERSION       AL672160.10 GI:25809448
KEYWORDS      HTG.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 192657)
REFERENCE
AUTHORS       Hammond, S.
TITLE         Direct Submission
JOURNAL
Submitted (17-APR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
On Nov 27, 2002 this sequence request: clonerequest@sanger.ac.uk
Sequence from the Mouse Genome Sequencing Consortium replaced gi:24394955.
shotgun may have been used to confirm this sequence. Whole genome
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations. Variation annotation may not be found in the sequence submission. This sequence was finished as described above, as we submit sequences with regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with the source databases:

```

FEATURES
SOURCE
1..192657
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-144K18"
/clone_1b="RP23-144K18"
/clone_1c="RP23-144K18"
BASS COUNT      51940 a 39496 c 42716 g 58505 t
ORIGIN
Query Match
Best Local Similarity 92.0%; Score 18.4; DB 10; Length 192657;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Db 156517 GGACAGCCCTGACAGCA 156536
1
|||||
1 GGACAGCCCTGACAGCA 20

```

```

RESULT 15
BX294664      193352 bp      DNA      linear      HTG 24-APR-2003
LOCUS         Mus musculus chromosome X clone RP23-384C22, *** SEQUENCING IN
DEFINITION
ACCESSION     BX294664
VERSION       BX294664.2 GI:29500936
KEYWORDS      HTG, HTGS, PHASE1, HTGS CANCELLED.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 193352)
REFERENCE
AUTHORS       Wclay, K.
TITLE         Direct Submission
JOURNAL
Submitted (23-APR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
On Apr 2, 2003 this sequence request: clonerequest@sanger.ac.uk
Sequence from the Mouse Genome Sequencing Consortium replaced gi:29134752.
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations. Variation annotation may not be found in the sequence submission. This sequence was finished as described above, as we submit sequences with regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with the source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/C.elegans/wormpep> RP23-144K18 is constructed by the group of Pletzer de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.

Location/Qualifiers

1..192657

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/chromosome="4"

/clone="RP23-144K18"

/clone\_1b="RP23-144K18"

/clone\_1c="RP23-144K18"

BASS COUNT 51940 a 39496 c 42716 g 58505 t

ORIGIN

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 10; Length 192657;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

\* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.

Sun Aug 3 09:03:40 2003

UB-09-834-291-10.rge

Search completed: August 1, 2003, 17:32:49  
Job time: 83.1573 secs

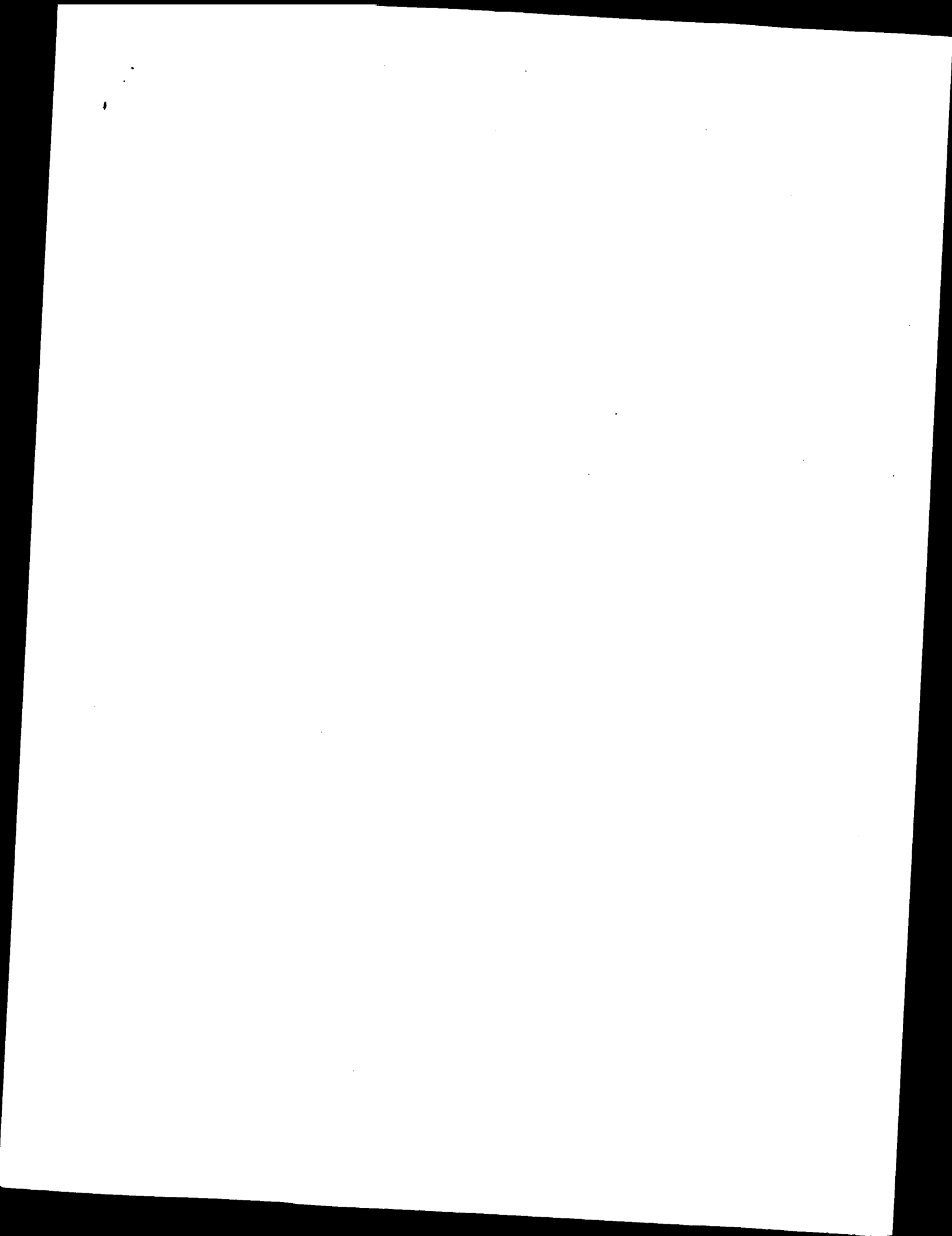
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 37461: contig of 37461 bp in length  
\* 37462 37561: gap of 100 bp  
\* 37562 45950: contig of 8389 bp in length  
\* 45951 46051: gap of 100 bp  
\* 46051 98426: contig of 52376 bp in length  
\* 98427 98527: gap of 100 bp  
\* 98527 118027: contig of 19501 bp in length  
\* 118028 118127: gap of 100 bp  
\* 118128 126520: contig of 8393 bp in length  
\* 126521 132811: contig of 100 bp  
\* 132812 132911: contig of 6191 bp in length  
\* 132912 172432: contig of 39521 bp in length  
\* 172433 172532: gap of 100 bp  
\* 172533 185735: contig of 13203 bp in length  
\* 185736 185835: gap of 100 bp  
\* 185836 190693: contig of 4858 bp in length  
\* 190694 190793: gap of 100 bp  
\* 190794 193352: contig of 2359 bp in length.  
Location/Qualifiers  
1. 193352  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="X"  
/clone="RP23-384C22"  
/clone\_lib="RPCT-23"  
1. 37461  
/note="assembly\_fragment:02244"  
fragment\_chain:1  
clone\_end:SP6  
vector\_side:left"  
37562. 45950  
/note="assembly\_fragment:00826"  
fragment\_chain:1"  
46051. 98426  
/note="assembly\_fragment:00297"  
fragment\_chain:1"  
98527. 118027  
/note="assembly\_fragment:01245"  
fragment\_chain:1"  
118128. 126520  
/note="assembly\_fragment:02356"  
fragment\_chain:1"  
126521. 132811  
/note="assembly\_fragment:01763"  
fragment\_chain:1"  
132912. 172432  
/note="assembly\_fragment:01130"  
fragment\_chain:1"  
172533. 185735  
/note="assembly\_fragment:01150"  
fragment\_chain:1"  
185836. 190693  
/note="assembly\_fragment:00830"  
fragment\_chain:1"  
clone\_end:T7  
vector\_side:right"  
190794. 193352  
/note="assembly\_fragment:00599"  
900 others

FEATURES  
source

misc\_feature  
1. 37461  
/note="assembly\_fragment:02244"  
fragment\_chain:1  
clone\_end:SP6  
vector\_side:left"  
37562. 45950  
/note="assembly\_fragment:00826"  
misc\_feature  
46051. 98426  
/note="assembly\_fragment:00297"  
misc\_feature  
98527. 118027  
/note="assembly\_fragment:01245"  
misc\_feature  
118128. 126520  
/note="assembly\_fragment:02356"  
misc\_feature  
126521. 132811  
/note="assembly\_fragment:01763"  
misc\_feature  
132912. 172432  
/note="assembly\_fragment:01130"  
misc\_feature  
172533. 185735  
/note="assembly\_fragment:01150"  
misc\_feature  
185836. 190693  
/note="assembly\_fragment:00830"  
misc\_feature  
190794. 193352  
/note="assembly\_fragment:00599"  
900 others  
BASE COUNT 58452 a 42616 c 39284 g 52100 t 900 others  
ORIGIN

Query Match 92.0%; Score 18.4; DB 2; Length 193352;  
Best Local Similarity 95.0%; Pred No. 68;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGNAGGCTGACAGCCA 20  
DB 36135 GGNAGGCTGACAGCCA 36116



Sun Aug 3 09:03:45 2003

US-09-834-291-10.FBT

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:13:22 ; Search time 68.4754 Seconds  
(without alignments)  
7098.748 Million cell updates/sec

US-09-834-291-10

Title: 1 gacacagcccgacacagca 20  
Sequence: 1 gacacagcccgacacagca 20Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthm:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hrc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hrc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estom:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vit:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rnd:\*

26: em\_gss\_vit:\*

27: em\_gss\_vit:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

## Description

Result No.	Score	Query Match Length	DB ID	Description
1	100.0	793	12	BT63679
2	100.0	962	9	AL540709
3	100.0	1089	13	BO072834
4	90.0	486	14	W45225

5	18	90.0	506	10	BG659530
6	18	90.0	753	13	BU142933
7	17.4	87.0	639	28	AZ939002
8	17.4	87.0	684	28	AZ450876
9	17.4	87.0	689	29	BX220214
10	17.4	87.0	701	29	CA605171
11	17.4	87.0	745	29	BX179961
12	17	85.0	429	28	AZ781046
13	17	85.0	531	9	AL195416
14	16.8	84.0	592	9	AM131236
15	16.8	84.0	629	28	AZ492025
16	16.8	84.0	730	28	BH898275
17	16.8	84.0	766	13	BU471804
18	16.8	84.0	843	14	CB560391
19	16.8	84.0	973	13	BO715536
20	16.8	84.0	1059	10	BF205338
21	16.8	84.0	1077	14	CD519542
22	16.8	84.0	1058	11	AK043415
23	16.8	84.0	1058	11	BF386638
24	16.4	82.0	191	10	AM131236
25	16.4	82.0	262	28	AQ632653
26	16.4	82.0	424	13	BY444129
27	16.4	82.0	545	10	BH303980
28	16.4	82.0	652	28	AG094812
29	16.4	82.0	652	28	AG094812
30	16.4	82.0	675	28	AG094812
31	16.4	82.0	696	10	BF966973
32	16.4	82.0	742	29	BZ328255
33	16.4	82.0	742	29	BG481356
34	16.4	82.0	802	10	BG296720
35	16.4	82.0	806	10	AG050636
36	16.4	82.0	809	28	CB192716
37	16.4	82.0	890	14	BU849141
38	16.4	82.0	926	13	BU849141
39	16.4	82.0	1114	29	CC278152
40	16.4	82.0	2193	10	BF579152
41	16.4	80.0	300	9	AL916663
42	16	80.0	406	14	CB807904
43	16	80.0	407	14	CB771883
44	16	80.0	417	14	CB766812
45	16	80.0	427	14	CB795385

## ALIGNMENTS

RESULT 1  
LOCUS BT63679/c 793 bp mRNA linear EST 25-SEP-2001  
DEFINITION 603049567F1 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5189752 5',  
mRNA sequence.

ACCESSION BT63679.1 GI:15755257

VERSION BT63679.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgsb@rockefeller.edu  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LIML at:  
http://image.llnl.gov  
Plate: LIML1474 row: e column: 17  
High quality sequence stop: 786.

## FEATURES

## Location/Qualifiers

1..962

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5189752"

/lab\_host="DH10B"

/clone\_id="NIH\_MGC\_116"

/note="Organ: pooled colon, kidney, stomach; Vector: PCMV-SPORT6; site: 1: NotI; site 2: EcoRV (destroyed); female, 71 yo male colon; 46 yo male kidney, and pool of 2 oligo-dt primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH MGC library."

## BASE COUNT

163 a 265 c 208 g 137 t

## Query Match

Best Local Similarity 100.0%; Score 20; DB 12; Length 793; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACAGCCCTGACAGCCA 20  
398 GGACAGCCCTGACAGCCA 379

## RESULT 2

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## source

## FEATURES

## Location/Qualifiers

1..962

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DE002YXN18"

/tissue\_type="PLACENTA"

/note="Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, the Not I and EcoRV sites of the PCMVSPORT 6 vector. Library was not normalized."

## FEATURES

## source

## BASE COUNT

280 a 263 c 333 g 212 t

## Query Match

Best Local Similarity 100.0%; Score 20; DB 13; Length 1089; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACAGCCCTGACAGCCA 20  
369 GGACAGCCCTGACAGCCA 388

## RESULT 4

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## source

## FEATURES

## Location/Qualifiers

1..1089

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5756324"

/tissue\_type="Leukocyte"

/lab\_host="DH10B"

/clone\_id="NIH\_MGC\_118"

/note="Vector: PCMV-SPORT6; Site 1: NotI; site 2: EcoRV (destroyed); RNA source: leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH MGC library."



```

Class: plasmid ends
High quality sequence stop: 619
Location/Qualifiers
    .. 619
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /strain="C57BL/6J"
    /db_xref="taxon:10090"

```

```

Class: Plasmid end
High quality sequence stop: 619.
Location/Qualifiers
    1. 619
        /organism="Mus musculus"
        /mol_type="genomic DNA"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="UUGC2M0197J18"
        /sex="Female"
        /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
        /clone_lib="Mouse 10kb plasmid UUGC2M library"
        /note="Vector: PWD29nv, Purified genomic DNA from M.
        musculus C57BL/6J (female) was obtained from the Jackson
        Laboratory Mouse DNA Resource
        (http://www.jax.org/resources/documents/dnares/). The DNA
        was hydridynamically sheared by repeated passage through a
        0.005 inch orifice at constant velocity. The sheared DNA
        was blunt end-repaired with T4 DNA polymerase and T4
        polynucleotide kinase. Adaptor oligonucleotides were
        ligated to the blunt ends in high molar excess. The
        10-15 kb range using prepreparative agarose gel
        electrophoresis. Vector DNA was prepared from a derivative
        of pMD42 (g14732114[gb|AF129072.1]) a copy-number
        inducible derivative of plasmid RI. The vector was ligated
        with adaptors complementary to the insert adaptors and
        adaptor vector DNA, and adaptor mouse DNA was annealed to
        chemically-competent E. coli XL10-Gold (Stratagene) cells
        and selected for ampicillin resistance."
    116 c
    87 g
    218 t

```

Query Match	87.0%	Score 17.4	DB 28	Length 619
Best Local Similarity	94.7%	Pred. No. 7.1e+02		
Matches 18	Conservative	0	Mismatches 1	Indels 0
Gaps 0				
1 GGACAAAGCCCTGGACAGCC	19			
595 GGACAAAGCTCTGGACAGCC	577			

RESULT 8	
AZ450876	
LOCUS	
DEFINITION	AZ450876 684 bp DNA linear GSS 04-OCT-2000
ACCESSION	U00249M17R Mouse 10kb plasmid U00CLM library Mus musculus genomic
VERSION	AZ450876
KEYWORDS	AZ450876.1 GI:10606116
SOURCE	GSS.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus
AUTHORS	Bukaryova; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Kelly, M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL	
COMMENT	Unpublished
Contact:	Robert B. Weiss

84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: [ddunn@genetics.utah.edu](mailto:ddunn@genetics.utah.edu)  
Insert Length: 10000 Std Error: 0.00  
Plate: 0249 row: M column: 17



## FEATURES

## SOURCE

Seq primer: CACACAGAAAACAGTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 684.  
 Location/Qualifiers

1..684  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUCG1M0249M17"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_lib="Mouse 10kb Plasmid UUCG1M library"  
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/nbases/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g1473214|3b|AF12072.1), a copy-number inducible derivative of plasmid RI. The vector was ligated with adaptor complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E.coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## BASE COUNT

180 a 202 c 160 g 142 t

## ORIGIN

Query Match 87.0%; Score 17.4; DB 28; Length 684;  
 Best Local Similarity 94.7%; Pred. No. 7, 3e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACAGCCCTGACAGCA 20  
 DB 611 GACAGCACTGACAGCA 629

RESULT 9  
 BX220214/c 689 bp DNA linear GSS 29-JAN-2003  
 LOCUS BX220214  
 DEFINITION Danio rerio genomic clone DKEX-269M11, genomic survey sequence.  
 ACCESSION BX220214  
 VERSION BX220214.1 GI:28052100  
 KEYWORDS GSS.  
 SOURCE Danio rerio (zebrafish)  
 ORGANISM Danio rerio  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.  
 1 (bases 1 to 689)  
 Humphray, S.J., Huckle, E. and Durham, J.L.  
 Direct Submision  
 Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humphray@sanger.ac.uk Unpublished  
 This sequence was generated from the sp6 end of BAC 269M11. 269M11 is part of the Daniolex BAC library created by R. Plaetzer and N.V. Keygene. Further details: http://www.sanger.ac.uk/projects/D\_rerio/  
 Location/Qualifiers

## FEATURES

## SOURCE

1..689  
 /organism="Danio rerio"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /clone="DKEX-269M11"  
 /tissue\_type="Testis"

BASE COUNT /note="vector pindigobAC-536"  
 180 a 144 c 144 g 221 t

## ORIGIN

Query Match 87.0%; Score 17.4; DB 29; Length 689;  
 Best Local Similarity 94.7%; Pred. No. 7, 4e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACAGCCCTGACAGCA 20  
 DB 419 GACAGCCCTGACAGCA 401

## RESULT 10

CA605171/c 701 bp mRNA linear EST 21-NOV-2002  
 LOCUS CA605171  
 DEFINITION w1.pK0046.e7 w1 Triticum aestivum cDNA clone w1.pK0046.e7 5' end  
 /mRNA sequence.

ACCESSION CA605171  
 VERSION CA605171.1 GI:25160333  
 KEYWORDS EST.  
 SOURCE Triticum aestivum (bread wheat)  
 ORGANISM Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae

REFERENCE 1 (bases 1 to 701)  
 Tingley, S.V., Powell, W., Wolters, P., Dolan, M., Halney, C., Yuan, Z., Miao, G., Caraher, N. and Hanley, W.K.  
 Dupont Wheat cDNA Sequence  
 Unpublished  
 CONTACT: Scott V. Tingley  
 Crop Genetics  
 E. I. Dupont de Nemours and Company  
 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA  
 Tel: 302-631-2602  
 Fax: 302-631-2607  
 Email: Scott.V.Tingley@USA.dupont.com  
 Seq primer: M3.

## FEATURES

## SOURCE

1..701  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4565"  
 /clone="w1.pK0046.e7"  
 /tissue\_type="root"  
 /clone\_lib="w1"  
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: XhoI; Wheat (Triticum aestivum L.) root; 7 day old seedling, light grown"

BASE COUNT 199 a 188 c 131 g 156 t 27 others  
 ORIGIN  
 Query Match 87.0%; Score 17.4; DB 14; Length 701;  
 Best Local Similarity 94.7%; Pred. No. 7, 4e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACAGCCCTGACAGCC 19  
 DB 255 GACAGCCCTGACAGCC 237

RESULT 11  
 BX179961/c 745 bp DNA linear GSS 28-JAN-2003  
 LOCUS BX179961  
 DEFINITION Danio rerio genomic clone DKEX-183M20, genomic survey sequence.  
 ACCESSION BX179961  
 VERSION BX179961.1 GI:28011764  
 KEYWORDS GSS.  
 SOURCE Danio rerio (zebrafish)  
 ORGANISM Danio rerio  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

REFERENCE 1  
AUTHORS Humphrey, S.J., Huckle, E. and Durham, J.L.  
TITLE Direct Submission  
JOURNAL Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humphrey@sanger.ac.uk Unpublished

COMMENT This sequence was generated from the SPE end of BAC 183M20. 183M20 is part of the Danilkey BAC library created by R. Plasterk and N.V. Keygene. Further details: [http://www.sanger.ac.uk/Projects/D\\_rerio/](http://www.sanger.ac.uk/Projects/D_rerio/).

FEATURES  
source Location/Qualifiers  
1..745

/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="DKEX-183M20"  
/tissue\_type="testis"  
/note="vector pindigobac-536"  
BASE COUNT 209 a 150 c 151 g 235 t

Query Match 87.0%; Score 17.4; DB 29; Length 745;  
Best Local Similarity 94.7%; Pred. No. 7.6e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACACGCTGACACGCA 20  
|||||  
419 GACACGCTGACACGCA 401

RESULT 12  
LOCUS T62311 331 bp mRNA linear EST 27-AUG-1998  
DEFINITION EGBST0074 TGRH Tachyzoite cDNA Toxoplasma gondii clone CGD008  
5' end, mRNA sequence.

ACCESSION T62311  
VERSION T62311.1 GI:665753  
KEYWORDS EST.

SOURCE  
ORGANISM Toxoplasma gondii  
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;  
Sarcocystidae; Toxoplasma.

REFERENCE 1 (bases 1 to 331)  
AUTHORS Wan, K.-L., Blackwell, J.M. and Ajioke, J.W.  
TITLE Toxoplasma gondii expressed sequence tags: insight into tachyzoite gene expression

JOURNAL Mol. Biochem. Parasitol. 75 (2), 179-186 (1995)  
MEDLINE 96258558  
PUBMED 8992316

COMMENT Contact: Ajioke JW  
Laboratory for Parasite Genome Analysis  
Cambridge University  
Department of Pathology, Tennis Court Road, Cambridge CB2 1QP, UK.  
Tel: 01223333923  
Fax: 01223333923  
Email: jajioke@gmp.mrc.ac.uk  
Clone information and library can be obtained from Jim Ajioke  
jwaemole.bio.cam.ac.uk  
Seq primer: SK.

FEATURES  
source Location/Qualifiers  
1..331

/organism="Toxoplasma gondii"  
/mol\_type="mRNA"  
/strain="RH"  
/db\_xref="taxon:5811"  
/clone="tgd008"  
/lab\_host="XLI-Blue MRF"  
/clone\_lib="TGRH Tachyzoite cDNA"  
/note="Vector: Lambda ZAP; Site 1: EcoRI; Site 2: XhoI;  
Toxoplasma RH strain tachyzoites were grown in human  
foreskin fibroblast cultures in vitro. The library was  
constructed by K.L. Wan, Cambridge University. CDNA's were

BASE COUNT 70 a 94 c 80 g 87 t  
synthesized from polyA RNAs by oligo d(T) priming and directionally cloned into the pCR1 to XhoI sites of the Lambda Zapit vector using the ZAP-cDNA synthesis kit (Stratagene). WARNING: the library contains a small percentage of cDNAs derived from the human host cells."

Query Match 85.0%; Score 17; DB 14; Length 331;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGGCTGACACG 17  
|||||  
314 GGACAGGCTGACACG 330

RESULT 13  
LOCUS A2781046 429 bp DNA linear GSS 16-FEB-2001  
DEFINITION 2M0018K23R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC2M0018K23 R, genomic survey sequence.

ACCESSION A2781046  
VERSION A2781046.1 GI:12913345  
KEYWORDS GSS.

SOURCE  
ORGANISM Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 429)  
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,  
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.  
and Wright, D.J. Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts

JOURNAL Unpublished  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: dunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00  
Plate: 0018 row: K column: 23  
Seq primer: CACACAGGAAACAGCTATGAC  
Class: plasmid ends

High quality sequence stop: 429.  
Location/Qualifiers  
1..429

/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC2M0018K23"  
/sex="Male"

/lab\_host="E. Coli strain XLI-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/note="Vector: pMD229v. Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(<http://www.jax.org/resources/documents/dnares/>). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adapored DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pMD42 (914732114|9b|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 125 a 88 c 126 g 90 t

Query Match 85.0%; Score 17; DB 28; Length 429;

Best Local Similarity 100.0%; Pred. No. 9.6e+02; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCTGACAG 17

Db 227 GGACAGCCTGACAG 243

RESULT 14 AA195416 531 bp mRNA linear EST 06-AUG-1997

LOCUS z36f10.s1 Soares\_NbHMPu\_S1 Homo sapiens cDNA clone IMAGE:665515

DEFINITION 3', mRNA sequence.

ACCESSION AA195416 GI:1785109

VERSION AA195416.1 GI:1785109

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 531) Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kucala, T., Lacy, M., Le, N., Lennon, G., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson R.

Washu-Merck EST Project 1997

Unpublished

CONTACT: Wilson R. Washington University School of Medicine

444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800 Fax: 314 286 1810

Email: est@wustl.wustl.edu

This clone is available royalty-free through LML; contact the

IMAGE Consortium (info@image.lml.gov) for further information.

Insert Length: 1020 Std Error: 0.00

High quality sequence stop: 436.

Location/Qualifiers

1..531

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/db\_xref="taxon:9606"

/clone="IMAGE:665515"

/tissue\_type="Pooled human melanocyte, fetal heart, and

pregnant uterus"

/lab\_host="DH10B"

/note="Organ: mixed (see below); Vector: pTTT3D-Pac

(Pharmacia) with a modified polylinker; Site 1: Not I;

Site 2: Eco RI; Equal amounts of plasmid DNA from three

normalized libraries (melanocyte 2NbHMP, pregnant uterus

NbHMP, and fetal heart NbH19W) were mixed, and ss circles

were made in vitro. Following HAP purification, this DNA

was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from pools of

5,000 clones made from the same 3 libraries. The pools

consisted of 1 M.A.G.E. clones 260232-265223,

340488-345479, and 484488-489479."

BASE COUNT 111 a 110 c 167 g 141 t 2 others

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 531;

Best Local Similarity 90.0%; Pred. No. 1.3e+03; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAGCCTGACAGCCA 20

Db 158 GGACAGCCTGACAGCCA 139

RESULT 15 AM11236 592 bp mRNA linear EST 27-OCT-1999

LOCUS x160c04.x1 NCI\_CGAP Gas4 Homo sapiens cDNA clone IMAGE:2622438 3'

DEFINITION similar to SW:TF1B HUMAN Q13263 TRANSCRIPTION INTERMEDIARY FACTOR

1-BETA; mRNA sequence.

ACCESSION AM11236 GI:6132843

VERSION AM11236.1 GI:6132843

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 592) Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

Tumor Gene Index

Unpublished

CONTACT: Robert Strausberg, Ph.D.

Email: cgapsb@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmett-Buck, M.D., Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LML at:

www.bio.lml.gov/bird/image/image.html

Seq primer: -40UP from G1bco

High quality sequence stop: 423.

Location/Qualifiers

1..592

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:2622438"

/tissue\_type="poorly differentiated adenocarcinoma with

signet ring cell features"

/lab\_host="DH10B"

/clone\_1b="NCI CGAP Gas4"

/note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.69 kb. Life Technologies catalog #:

11549-011"

BASE COUNT 133 a 146 c 194 g 115 t 4 others

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 592;

Best Local Similarity 90.0%; Pred. No. 1.3e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAGCCTGACAGCCA 20

Db 590 GGACAGCCTGACAGCCA 571

Search completed: August 1, 2003, 20:42:58

Job time : 72.4754 secs



Sun Aug 3 09:03:41 2003

us-09-834-291-10.rng

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:32:17 ; Search time 6.69868 Seconds  
(without alignments)  
8059.612 Million cell updates/sec

Title: US-09-834-291-10  
Perfect score: 1 ggcacagccctgacagcca 20  
Sequence: 1 ggcacagccctgacagcca 20

Scoring table: IDENTITY: NUC  
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues  
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: N\_Geneseq\_19jun03:\*  
2: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1980.DAT:\*  
3: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1981.DAT:\*  
4: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1982.DAT:\*  
5: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1983.DAT:\*  
6: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1984.DAT:\*  
7: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1985.DAT:\*  
8: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1986.DAT:\*  
9: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1987.DAT:\*  
10: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1988.DAT:\*  
11: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1989.DAT:\*  
12: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1990.DAT:\*  
13: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1991.DAT:\*  
14: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1992.DAT:\*  
15: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1993.DAT:\*  
16: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1994.DAT:\*  
17: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1995.DAT:\*  
18: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1996.DAT:\*  
19: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1997.DAT:\*  
20: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1998.DAT:\*  
21: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA1999.DAT:\*  
22: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA2000.DAT:\*  
23: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA2001.DAT:\*  
24: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA2002.DAT:\*  
25: /SIDS1/gcgdata/geneSeq/geneSeq-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	266	AAZ88700	Human CD95 receptor
2	16.4	82.0	516	ABX1757	Murine gene trappe
3	16.4	82.0	8243	ABZ74553	Secreted protein g
4	16.4	82.0	12710	AAK81012	Human immune/haema
5	16.4	82.0	12710	ABZ74552	Secreted protein g
6	16	80.0	433	ABX52814	Bovine ESR associa
7	15.8	79.0	165	ABX27459	Human GDP-nanose
8	15.8	79.0	261	ABZ41531	Human gonorrhoea nuc

C	9	15.8	79.0	273	20	AAZ1974	Neisseria meningit
C	10	15.8	79.0	288	21	AAZ53178	Neisseria meningit
C	11	15.8	79.0	309	20	AAZ53176	Neisseria gonorrho
C	12	15.8	79.0	381	20	AAZ11975	Neisseria gonorrho
C	13	15.8	79.0	409	22	AAK58992	Human immune/haema
C	14	15.8	79.0	749	22	AAK74242	Human immune/haema
C	15	15.8	79.0	749	22	AAK74243	Human immune/haema
C	16	15.8	79.0	750	22	AAK74241	Human immune/haema
C	17	15.8	79.0	1012	23	AAZ59679	Protonibacterium
C	18	15.8	79.0	1482	25	ACA01448	C. glutamicum deri
C	19	15.8	79.0	1482	25	ACA01448	C. glutamicum deri
C	20	15.8	79.0	16949	23	AAZ59585	Protonibacterium
C	21	15.8	79.0	39325	22	AAZ59585	Human immune/haema
C	22	15.8	79.0	88421	24	AAZ40781	8842int genomic DN
C	23	15.8	79.0	136328	24	AAZ35015	Human gene express
C	24	15.8	79.0	155074	24	AAZ85735	Human genomic regi
C	25	15.8	79.0	309400	22	AAZ68534	C glutamicum codin
C	26	15.4	77.0	272	25	ABX54046	Bovine ESR associa
C	27	15.4	77.0	401	22	AAZ95694	Human neutregulin g
C	28	15.4	77.0	401	22	AAZ97187	Human neutregulin g
C	29	15.4	77.0	401	24	ABT01957	Human neutregulin-1
C	30	15.4	77.0	498	24	ABN68645	Human neutregulin-1
C	31	15.4	77.0	673	22	AAZ35624	Streptococcus poly
C	32	15.4	77.0	673	25	ABX58612	Human musculoskele
C	33	15.4	77.0	758	20	AAZ16661	CDNA encoding nove
C	34	15.4	77.0	845	24	AAZ62770	Human musculoskele
C	35	15.4	77.0	860	22	AAZ7561	Human musculoskele
C	36	15.4	77.0	860	22	ABX60548	CDNA encoding nove
C	37	15.4	77.0	860	25	ABX60549	CDNA encoding nove
C	38	15.4	77.0	921	21	AAZ93516	Human secreted pro
C	39	15.4	77.0	921	21	AAZ93516	Human secreted pro
C	40	15.4	77.0	1560	21	AAZ97901	Human secreted pro
C	41	15.4	77.0	2010	21	AAZ59604	Human secreted pro
C	42	15.4	77.0	2059	21	AAZ59970	Human secreted pro
C	43	15.4	77.0	3038	25	ABT19900	Aspergillus fumiga
C	44	15.4	77.0	3604	25	ABT19714	Aspergillus fumiga

ALIGNMENTS

AAZ88700	AAZ88700 standard; DNA; 266 BP.
AAZ88700	AAZ88700; (first entry)
AAZ88700	11-MAY-2000 (first entry)
AAZ88700	Human CD95 receptor intron 1 fragment.
AAZ88700	p53; CD95 receptor; human; screening; apoptosis-modulation;
AAZ88700	Cancer chemotherapy; ss.
AAZ88700	Homo sapiens.
AAZ88700	Key
AAZ88700	Location/Qualifiers
AAZ88700	protein_bind
AAZ88700	160..179
AAZ88700	/tag= a
AAZ88700	/bound_molecy= p53
AAZ88700	DE19847779-Cl.
AAZ88700	03-FEB-2000.
AAZ88700	16-OCT-1998; 98DE-1047779.
AAZ88700	16-OCT-1998; 98DE-1047779.
AAZ88700	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
AAZ88700	Krammer P, Mueller-Schilling M, Oren M;

XX WPI; 2000-162245/15.

XX Novel receptor DNA useful for identifying apoptosis-modulating  
XX substances potentially useful for cancer chemotherapy -

PS Claim 2; Fig 4; 12pp; German.

CC This invention describes a novel p53-binding region of a human CP95  
CC receptor DNA molecule. The p53-binding region, or a vector containing  
CC it, can be used to screen for apoptosis-modulating substances  
CC potentially useful for cancer chemotherapy. This sequence represents a  
CC fragment of the human CP95 receptor intron 1 which contains a p53 binding  
CC region described in the method of the invention.

SO Sequence 266 BP; 49 A; 72 C; 110 G; 35 T; 0 other;

Query Match Best Local Similarity 100.0%; Score 20; DB 21; Length 266;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGACAGCCCTGACAGCCA 20  
DB 160 GGACAGCCCTGACAGCCA 179

RESULT 2  
ABX91757/C

ID ABX91757 standard; cDNA; 516 BP.

AC ABX91757;

DT 07-MAY-2003 (first entry)

DE Murine gene trapped sequence (GTS) SEQ ID NO 1101.

XX Murine; mouse; gene trap technology; gene trapped sequence; GTS;  
KW gene identification; functional genomic analysis; gene discovery;  
KW gene expression analysis; cross species hybridisation analysis;  
KW antisense inhibition; gene targeting; gene; ss.

OS Mus sp.

PN US2002161207-A1.

PD 31-OCT-2002.

PF 30-NOV-2000; 2000US-0728444.

PR 01-DEC-1999; 99US-168360P.

PA (FRIE// FRIEDRICH G.  
PA (ZAMB// ZAMBROWICZ B.  
PA (SAND// SANDS A T.

PI Friedrich G, Zambrowicz B, Sands AT;  
PI WPI; 2003-288124/28.

PT New murine polynucleotides comprising gene trapped sequences, useful in  
PT functional genomic analysis, in the development of new therapeutic or  
PT diagnostic agents, for diagnostic gene expression analysis or for  
PT genetic manipulations -

PS Claim 2; SEQ ID NO 1101; 29pp; English.

CC The present invention relates to novel murine cDNAs produced using  
CC gene trap technology. The OMNIBANK gene trapped sequences (GTSs)  
CC are individually identified novel genes, and are useful in functional  
CC genomic analysis, in the discovery and development of new therapeutic  
CC and diagnostic agents, for gene discovery, for diagnostic gene  
CC expression analysis, for cross species hybridisation analysis, and for  
CC genetic manipulations such as antisense inhibition or gene targeting.

CC The polynucleotides of the invention are also useful for isolating  
CC cDNAs, genomic clones or full-length genes/polynucleotides, or their  
CC homologues, heterologues, paralogues or orthologues, that are capable  
CC of hybridising to one or more of the new murine polynucleotide  
CC sequences. The polynucleotides are also useful for identifying the  
CC coding regions of the murine genome, and as hybridisation probes.  
CC ABX30657-ABX91862 represent the murine GTSs of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC USPTO web site at [seqdata.uspto.gov/psipdsidentry.html](http://seqdata.uspto.gov/psipdsidentry.html).

SO Sequence 516 BP; 163 A; 90 C; 118 G; 140 T; 5 other;

Query Match Best Local Similarity 82.0%; Score 16.4; DB 25; Length 516;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GACAGCCCTGACAGCC 19  
DB 100 GACAGCCCTGACAGCC 83

RESULT 3  
ABZ74553

ID ABZ74553 standard; DNA; 8243 BP.

AC ABZ74553;

DT 12-MAY-2003 (first entry)

DE Secreted protein gene 346 genomic fragment HTOH021, SEQ ID NO:1700.

XX Human; secreted protein; cancer; tumour; hyperproliferative disorder;  
KW autoimmune disorder; inflammation; angiogenic diseases; AIDS; drug  
KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;  
KW cytotoxic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;  
KW antineoplastic; vulnery; chromosome 15q24-25; gene; ds.

OS Homo sapiens.

PN WO200277013-A2.

PD 03-OCT-2002.

PF 26-MAR-2002; 2002WO-US09370.

PR 27-MAR-2001; 2001US-278650P.

PR 12-SEP-2001; 2001US-0950082.

PR 12-SEP-2001; 2001US-0950083.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

PI WPI; 2003-040578/03.

PT New human secreted proteins and nucleic acids, useful for detecting or  
PT treating cancer or other hyperproliferative disorders, autoimmune  
PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia -  
PS Disclosure; Page 2282-2284; 2474pp; English.

CC ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted  
CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.  
CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The  
CC invention also encompasses antibodies specific for the secreted proteins,  
CC the use of the secreted proteins in drug screening and recombinant  
CC vectors and host cells comprising a nucleic acid of the invention. The  
CC associated proteins are thought to be involved in biological activities  
CC associated with cellular signalling, cellular differentiation, cell  
CC migration, prothrombin activation and neurotransmitter activity. The  
CC secreted proteins, nucleic acids encoding them, antibodies or antibody

fragments specific for the secreted proteins, and modulators of protein activity are useful for diagnosing or treating cancers or other hyperproliferative disorders. Additionally, the secreted proteins and their nucleic acids may also be used in the treatment of autoimmune disorders, inflammatory disorders, diseases involving angiogenesis, AIDS (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote wound healing. Nucleic acids of the invention may be used for chromosome identification, chromosome mapping, in gene therapy, for identifying individuals from minute biological samples, as hybridisation probes, and as molecular weight markers. The present sequence represents a human secreted protein genomic fragment referred to in the disclosure of the invention.

Sequence 8243 BP; 2717 A; 1618 C; 1711 G; 2197 T; 0 other;

Query Match 82.0%; Score 16.4; DB 25; Length 8243;  
Best Local Similarity 94.4%; Pred. No. 1.9e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACAGCCCTGACAGCC 19  
DB 8148 GACAGCCCTGACAGCC 8165

RESULT 4  
AAK81012/C  
ID AAK81012 standard; DNA; 12710 BP.

AC AAK81012;  
DT 07-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35824.  
XX Human immune/haematopoietic; immune/haematopoietic antigen; cancer;  
KM cytoablastic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX MO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-020515.

XX 19-MAY-2000; 2000US-0209467.

XX 07-JUN-2000; 2000US-0214886.

XX 28-JUN-2000; 2000US-0215135.

XX 30-JUN-2000; 2000US-021647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

XX 14-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 26-JUL-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226868.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.

PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0232397.

PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 14-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.

PR 25-SEP-2000; 2000US-0234998.

PR 26-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.

PR 27-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.

PR 29-SEP-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241809.

PR 20-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.

PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.

PR 08-NOV-2000; 2000US-0246609.

PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	01-DEC-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	05-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251968.
PR	06-DEC-2000;	2000US-0256719.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251889.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259578.

(HOMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -

Disclosure; SEQ ID NO 35824; 3071pp + Sequence listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAK68170 to AAK691921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. (I) example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent the diagnosis and treat immune/haematopoietic-related diseases, especially to AAK68794 represent metastases of haematopoietic-derived cells. AAK64703 sequences from the present invention. AAK54942 to AAK54950 and AAK62169 represent sequences used in the exemplification of the present invention.

Sequence 12710 BP; 3496 A. 3577 C

Query Match	Score	DB	Length
Test Local Similarity	82.0%	16.4	22
atches 17; Conservative	94.4%		
	0	Mismatches	1
		Indels	0
		Gaps	0

Db

96 GACAGCCCTGAGAACC 79

RESULT 5  
ABZ74552  
ID ABZ74552 standard; DNA; 12710 BP.  
XX

12-MAY-2003 (first entry)

Secreted protein gene 346 genomic fragment HTOHO21, SEQ ID NO:1699

autoimmunity; cancer; tumour; hyperproliferative disorder;  
 acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing  
 cytostatic; gene therapy; antiinflammation; chromosome mapping;  
 antineoplastic; vulnery; chromosome 15q24-25; gene; ds.  
 Homo sapiens

WO200277013-A2.

03-OCT-2002.

2002-2003; 2002ZWO-US093370.

2-SEP-2001; 2001US-278650P.  
12-SEP-2001; 2001US-0950082.  
12-SEP-2001; 2001US-0950083.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;

WPI; 2003-040578/03.

New human secreted p

treating cancer or other hyperproliferative disorders, inflammatory disorders, autoimmune disorders, and nucleic acids, useful for detecting or secreting proteins and nucleic acids.

Disclosure; Page 2278-2281; 2474pp; English.

ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted protein genes, and ABP00947-ABP01353 represent the proteins they encode. ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The invention also encompasses antibodies specific for the secreted proteins. The use of the secreted proteins in drug screening and recombinant vectors and host cells comprising a nucleic acid of the invention, associated proteins are thought to be involved in biological activities migration, prohomone activation and neurotransmitter activity. Cell secreted proteins, nucleic acids encoding them, antibodies or antibody fragments specific for the secreted proteins, and modulators of protein activity are useful for diagnosing or treating cancers or other hyperproliferative disorders. Additionally, the secreted proteins and their nucleic acids may also be used in the treatment of autoimmune disorders, inflammatory disorders, diseases involving angiogenesis, AIDS (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote wound healing. Nucleic acids of the invention may be used for chromosome identification, chromosome mapping, in gene therapy, for identifying individuals from minute biological samples, as hybridisation probes, as molecular weight markers. The present sequence represents a human secreted protein genomic fragment referred to in the disclosure of the invention.

```
Sequence 12710 BP; 4072 A; 2575 C; 2567 G; 3496 T; 0 other;
every Match      82.0%; Score 16.4; DB 25; Length 12710;
Self Local Similarity 94.4%; Pred. No. 2+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0
```



QY 2 GACAGCCTGACAGCC 19  
 |||||  
 DB 12615 GACAGCCTGACAGCC 12632

RESULT 6  
 ABX52814/C  
 ID ABX52814 standard; cDNA; 433 BP.

XX AC ABX52814;

DT 25-FEB-2003 (first entry)

XX DE Bovine EST associated with lactation/muscle/fat deposition #2743.

XX DE Bovine; sex; EST; expressed sequence tag; lactation; LMPD;

KW muscle deposition; fat deposition; genome mapping; gene identification;

KW gene analysis; cattle breeding.

XX OS Bos Taurus.

XX PN US2002137160-A1.

XX PD 26-SEP-2002.

XX PF 26-OCT-2001; 2001US-0983965.

XX PR 17-DEC-1998; 98US-113678P.

XX PR 15-DEC-1999; 99US-0465231.

XX PA (BYAT/) BYATT J C.

XX PA (MATH/) MATHIALAGAN N.

XX PA (TAON/) TAO N.

XX PA (WARR/) WARREN W C.

XX PI Byatt JC, Mathialagan N, Tao N, Warren WC;

XX WPI; 2003-102386/09.

XX PT Purified nucleic acid molecules, useful for genome mapping, gene

XX PT identification and analysis, cattle breeding or preparation of

XX PT constructs for cattle gene expression and genetically improved cattle -

XX PS Claim 2; SEQ ID NO 2743; 38pp; English.

XX PS The invention relates to a purified nucleic acid molecule associated with

XX CC lactation or muscle and fat deposition (designated LMPD), derived to a

XX CC from cattle, and the LMPD nucleic acid can specifically hybridize to a

XX CC second nucleic acid molecule comprising any of 5912 nucleotide

XX CC sequences, appearing as ABX50072-ABX55983, or complements of them.

XX CC Also included are: (1) a transformed cell having a nucleic acid

XX CC comprising an LMPD nucleic acid linked to a promoter and a 3' end

XX CC transcription and addition of polyadenylated ribonucleotides to a 3' end

XX CC of the mRNA molecule, and (2) determining a level or pattern of a

XX CC nucleic acid (comprising any of the 5912 nucleic acid sequences or its

XX CC complement or fragment) with a complementary nucleic acid molecule

XX CC obtained from the bovine cell or tissue, where hybridization between the

XX CC marker nucleic acid and the complementary nucleic acid permits the

XX CC detection of the molecule; and (b) detecting the level or pattern of the

XX CC complementary nucleic acid, where the detection of the complementary

XX CC nucleic acid is predictive of the level or pattern of the molecule.

XX CC The LMPD nucleic acid is used for determining a level or pattern

XX CC of a molecule in a bovine cell or tissue. It is useful for genome

XX CC mapping, gene identification and analysis, cattle breeding, preparation

XX CC of constructs for use in cattle gene expression, or for genetically

XX SQ Sequence 433 BP; 86 A; 88 C; 159 G; 100 T; 0 other;  
 Query Match 80.0%; Score 16; DB 25; Length 433;  
 Best Local Similarity 100.0%; Pred. No. 2e+02; Indels 0; Gaps 0;  
 Matches 16; Conservative 0; Mismatches 0

QY 4 CAAGCCTGACAGCC 19  
 |||||  
 DB 102 CAAGCCTGACAGCC 87

RESULT 7  
 ABX27459  
 ID ABX27459 standard; cDNA; 165 BP.

XX AC ABX27459;

DT 11-FEB-2003 (first entry)

XX DE Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #9516.

XX DE Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; sex; inflammation;

XX KW cellular fucosylation; glycoconjugate fucosylation; transplant rejection;

XX KW arthritis; asthma; sepsis; reperfusion injury; stroke; infection;

XX KW complex carbohydrate; gene replacement therapy; immunosuppressive;

XX KW antiinflammatory; antiarthritic; antibacterial; cerebroprotective;

XX KW antisthmatic; vasotropic.

XX OS Homo sapiens.

XX PN US2002110548-A1.

XX PD 15-AUG-2002.

XX PF 11-JUN-2001; 2001US-0878574.

XX PR 22-NOV-1996; 96US-0753233.

XX PR 03-DEC-1997; 97US-0984246.

XX PR 09-SEP-1998; 98US-0149674.

XX PR 14-JUN-1999; 99US-0333177.

XX PA (GENW) GENENTICS INST INC.

XX PI Sullivan F, Kriz R, Kumar R;

XX WPI; 2003-066673/06.

XX PT New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)

XX PT peptide, for manufacturing complex carbohydrates, or as targets for

XX PT screening GM4,6D antagonists for treating e.g. arthritis, or transplant

XX PS Disclosure; SEQ ID NO 9518; 6pp; English.

XX PS The invention relates to a composition comprising a human GDP-mannose

XX CC 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying

XX CC GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation

XX CC in a mammalian subject and for treating or ameliorating diseases affected

XX CC by the level of cellular fucosylation or diseases include arthritis,

XX CC fucosylation of glycoconjugates. These diseases include stroke or

XX CC transplant rejection, asthma, sepsis, reperfusion injury, stroke or

XX CC infection. The GM4,6D peptide, or a polynucleotide encoding it is also

XX CC useful for manufacturing complex carbohydrates and as targets for

XX CC screening small molecule antagonists of the activity of the enzyme. The

XX CC polynucleotide is useful in developing an assay for defects in the

XX CC enzyme, as well as in gene replacement therapy. Sequences

XX CC ABX1794-ABX1794 and ABX1794-ABX33716 represent DNA molecules encoding

XX CC human GM4,6D peptides of the invention.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC at seqdata.uspto.gov/sequence.html.

SO Sequence 165 BP; 34 A; 66 C; 42 G; 23 T; 0 other;

Query Match 79.0%; Score 15.8; DB 25; Length 165;  
Best Local Similarity 89.5%; Pred. No. 2.3e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCACAGCCCTGACAGCC 19  
Db 131 GCACAGCCCTGACAGCC 149

## RESULT 8

AB241531  
ID AB241531 standard; DNA; 261 BP.

AC AB241531;

DT 07-MAR-2003 (first entry)

DE N. gonorrhoeae nucleotide sequence SEQ ID 7651.

KW Antibacterial; infection; vaccine; gene therapy; gene; ds.

OS Neisseria gonorrhoeae.

PN MO200279243-A2.

PD 10-OCT-2002.

PR 12-FEB-2002; 2002WO-IB02069.

PR 12-FEB-2001; 2001GB-0003424.

PA (CHIR-) CHIRON SPA.

PI Fontana MR, Pizsa M, Massignani V, Monaci E;

DR WPI; 2003-058415/05.

DR F-PSDB; ABP80561.

PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a  
medicament for treating or preventing N. gonorrhoeae infection -  
PS Disclosure; Page 746; 815pp; English.

CC The present invention relates to proteins from Neisseria gonorrhoeae.  
CC Antibodies that specifically bind to the proteins encoding the proteins and  
CC comprising the protein, nucleic acid or antibody is useful for the  
CC manufacture of a medicament for treating or preventing N. gonorrhoeae  
CC infection, this may be in the form of a vaccine or gene therapy.  
CC Sequences given in records AB237706-AB242016 represent nucleic acid  
CC molecules of the invention.

SO Sequence 261 BP; 53 A; 71 C; 65 G; 72 T; 0 other;

Query Match 79.0%; Score 15.8; DB 25; Length 261;  
Best Local Similarity 89.5%; Pred. No. 2.4e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GACAGCCCTGACAGCCA 20  
Db 153 GACAGCCCTGACAGCCA 171

## RESULT 9

AA211974/c

ID AA211974 standard; DNA; 273 BP.

AC AA211974;

DT 08-OCT-1999 (first entry)

DE Neisseria meningitidis strain A complete ORF37 sequence.

KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

OS treatment; Neisseria infection; meningitis; septicemia; gonorrhea; ss.

XX Neisseria meningitidis.

PN MO9924578-A2.

PD 20-MAY-1999.

PR 09-OCT-1998; 98WO-IB01665.

PR 01-SEP-1998; 98GB-0019016.

PR 06-NOV-1997; 97GB-0023516.

PR 14-NOV-1997; 97GB-0024190.

PR 18-NOV-1997; 97GB-0024386.

PR 27-NOV-1997; 97GB-0025158.

PR 10-DEC-1997; 97GB-0026147.

PR 14-JAN-1998; 98GB-0000759.

(CHIR-) CHIRON SPA.

PI Grandi G, Massignani V, Pizsa M, Rappuoli R, Scarlato V;

DR WPI; 1999-327407/27.

DR P-PSDB; AAY38501.

PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for  
diagnosis, treatment and prevention of infection  
PS Claim 3; Page 61; 524pp; English.

CC Nucleotide sequences AA211972-212358 represent open reading frames  
CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode  
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their  
CC fragments, their nucleic acids and antibodies are used for diagnosis,  
CC prevention (as vaccines) or treatment of Neisseria infections,  
CC such as meningitis, septicemia and gonorrhea. Both organisms  
CC are closely related. Fragments of the nucleic acids are useful  
CC as hybridisation probes and antisense reagents.

SO Sequence 273 BP; 75 A; 67 C; 83 G; 48 T; 0 other;

Query Match 79.0%; Score 15.8; DB 20; Length 273;  
Best Local Similarity 89.5%; Pred. No. 2.4e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GACAGCCCTGACAGCCA 20  
Db 217 GACAGCCCTGACAGCCA 199

## RESULT 10

AA253178

ID AA253178 standard; DNA; 288 BP.

AC AA253178;

DT 21-MAR-2000 (first entry)

DE Neisseria meningitidis ORF 091 partial DNA sequence SEQ ID NO:307.

KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

OS antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;

XX antibacterial; gene therapy; ds.

OS Neisseria meningitidis.

PN MO9957280-A2.

PD 11-NOV-1999.

PR 30-APR-1999; 99WO-US09346.  
 PR 01-MAY-1998; 98US-0083758.  
 PR 31-JUL-1998; 98US-0094869.  
 PR 02-SEP-1998; 98US-0098994.  
 PR 02-SEP-1998; 98US-0099062.  
 PR 09-OCT-1998; 98US-0103749.  
 PR 09-OCT-1998; 98US-0103794.  
 PR 09-OCT-1998; 98US-0103796.  
 PR 25-FEB-1999; 99US-0121528.  
 PA (CHIR) CHIRON CORP.  
 PA (GENO-) INST GENOMIC RES.  
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
 PI Tettelin H, Venter JC;  
 PI WPI: 2000-062150/05.  
 DR P-PSDB; AAY74416.  
 DR Novel Neisserial polypeptides predicted to be useful antigens for  
 PT vaccines and diagnostics -  
 PS Claim 7; Page 287; 1453pp; English.  
 XX AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941  
 CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides  
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent  
 CC PCR primers used in the exemplification of the present invention. The  
 CC polypeptides, the polynucleotides, antibodies and compositions of  
 CC the invention can be used as vaccines, as diagnostic reagents, and as  
 CC immunogenic compositions. The polypeptides can be used in the  
 CC manufacture of medicaments for treating or preventing infection due to  
 CC *Neisseria meningitidis* (e.g. meningitis and septicemia), to detect the  
 CC presence of *Neisseria meningitidis*, or to raise antibodies. They may also  
 CC be used to screen for agonists or antagonists, which may themselves  
 CC have use as antibacterial agents. The polynucleotides of the invention  
 CC may also be used in gene therapy protocols.  
 CC Sequence 288 BP; 57 A; 73 C; 74 G; 84 T; 0 other;  
 SQ  
 Query Match 79.0%; Score 15.8; DB 21; Length 288;  
 Best Local Similarity 89.5%; Pred. No. 2.4e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GACAAAGCCCTGACCAAGCCA 20  
 DB 198 GACAAAGCCCTGACCAAGCCA 216  
 RESULT 11  
 AA253176  
 ID AA253176 standard; DNA; 309 BP.  
 AC AA253176;  
 XX  
 DT 21-MAR-2000 (first entry)  
 XX  
 DE *Neisseria gonorrhoeae* ORF 091 partial DNA sequence SEQ ID NO:303.  
 XX  
 KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;  
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
 KW antibacterial; gene therapy; ds.  
 XX  
 OS *Neisseria gonorrhoeae*.  
 XX  
 PN WO9957280-A2.  
 XX  
 PD 11-NOV-1999.  
 XX  
 PF 30-APR-1999; 99WO-US09346.  
 XX

PR 01-MAY-1998; 98US-0083758.  
 PR 31-JUL-1998; 98US-0094869.  
 PR 02-SEP-1998; 98US-0098994.  
 PR 02-SEP-1998; 98US-0099062.  
 PR 09-OCT-1998; 98US-0103749.  
 PR 09-OCT-1998; 98US-0103794.  
 PR 09-OCT-1998; 98US-0103796.  
 PR 25-FEB-1999; 99US-0121528.  
 PA (CHIR) CHIRON CORP.  
 PA (GENO-) INST GENOMIC RES.  
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
 PI Tettelin H, Venter JC;  
 PI WPI: 2000-062150/05.  
 DR P-PSDB; AAY74416.  
 DR Novel Neisserial polypeptides predicted to be useful antigens for  
 PT vaccines and diagnostics -  
 PS Claim 7; Page 287; 1453pp; English.  
 XX AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941  
 CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides  
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent  
 CC PCR primers used in the exemplification of the present invention. The  
 CC polypeptides, the polynucleotides, antibodies and compositions of  
 CC the invention can be used as vaccines, as diagnostic reagents, and as  
 CC immunogenic compositions. The polypeptides can be used in the  
 CC manufacture of medicaments for treating or preventing infection due to  
 CC *Neisseria meningitidis* (e.g. meningitis and septicemia), to detect the  
 CC presence of *Neisseria meningitidis*, or to raise antibodies. They may also  
 CC be used to screen for agonists or antagonists, which may themselves  
 CC have use as antibacterial agents. The polynucleotides of the invention  
 CC may also be used in gene therapy protocols.  
 CC Sequence 309 BP; 64 A; 82 C; 78 G; 85 T; 0 other;  
 SQ  
 Query Match 79.0%; Score 15.8; DB 21; Length 309;  
 Best Local Similarity 89.5%; Pred. No. 2.4e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GACAAAGCCCTGACCAAGCCA 20  
 DB 198 GACAAAGCCCTGACCAAGCCA 216  
 RESULT 12  
 AA211975/C  
 ID AA211975 standard; DNA; 361 BP.  
 AC AA211975;  
 XX  
 DT 08-OCT-1999 (first entry)  
 XX  
 DE *Neisseria gonorrhoeae* complete ORF37 sequence.  
 XX  
 KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;  
 KW treatment; *Neisseria* infection; meningitis; septicemia; gonorrhea; ss.  
 XX  
 OS *Neisseria gonorrhoeae*.  
 XX  
 PN WO9924578-A2.  
 XX  
 PD 20-MAY-1999.  
 XX  
 PF 09-OCT-1998; 98WO-IB01665.  
 XX  
 PD 01-SEP-1998; 98GB-0019016.  
 PR 06-NOV-1997; 97GB-0023516.  
 PR 14-NOV-1997; 97GB-0024190.  
 PR

PR 18-NOV-1997; 97GB-0024386.  
PR 27-NOV-1997; 97GB-0025158.  
PR 10-DEC-1997; 97GB-0026147.  
PR 14-JAN-1998; 98GB-0000759.

PA (CHIR-) CHIRON SPA.

PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;  
DR WPI, 1999-327407/27.

DR P-PSDB; AAY38502.

PT Proteins from *Neisseria meningitidis* and *N. gonorrhoeae* useful for  
XX diagnosis, treatment and prevention of infection

PS Claim 3; Page 61; 524pp; English.

CC Nucleotide sequences AA211972-212358 represent open reading frames  
CC (ORFs) of *Neisseria meningitidis* and *N. gonorrhoeae* which encode  
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their  
CC fragments, their nucleic acids and antibodies are used for diagnosis, their  
CC prevention (as vaccines) or treatment of *Neisseria* infections,  
CC such as meningitis, septicemia and gonorrhea. Both organisms,  
CC as are closely related. Fragments of the nucleic acids are useful  
CC as hybridisation probes and antisense reagents.

XX Sequence 381 BP; 106 A; 84 C; 114 G; 77 T; 0 other;

Query Match 79.0%; Score 15.8; DB 20; Length 381;  
Best Local Similarity 89.5%; Pred. No. 2.5e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Dn 2 GACAGCCTTGACAGCA 20  
325 GACAGCCTTGACAGCA 307

RESULT 13

AAKS8992  
ID AAKS8992 standard; cDNA; 409 BP.

XX AAKS8992;

DT 06-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:4052.

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytosolic; gene therapy; vaccine; metastasis; ss.

XX Homo sapiens.

PN W0200157182-A2.

PD 09-AUG-2001.

PF 17-JAN-2001; 2001WO-US01354.

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 11-JUL-2000; 2000US-0216860.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 14-AUG-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225256.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0225759.

PR 22-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226686.

PR 23-AUG-2000; 2000US-0227182.

PR 30-AUG-2000; 2000US-0227009.

PR 01-SEP-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 05-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 06-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 08-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 12-SEP-2000; 2000US-0232081.

PR 14-SEP-2000; 2000US-0232397.

PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 21-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 25-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.

PR 26-SEP-2000; 2000US-0234998.

PR 27-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.

PR 02-OCT-2000; 2000US-0236370.

PR 02-OCT-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.

PR 13-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239335.

PR 20-OCT-2000; 2000US-0239337.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.

PR 01-NOV-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.



(AUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM,  
WPI; 2001-483426/52.

Sequence 749 bp, 301 nt, 100% identity to AAK54992 (GenBank accession number AF022509.1) and 100% identity to AAK54991 (GenBank accession number AF022508.1). The amino acid sequences given in AAM82170 to AAM91921. (1) have cytostatic activity, and can be used in gene therapy and vaccine production. (1) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (1) expression. (1) example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (1) by expressing inactive proteins or to supplement the patient's own production of (1). Additionally, (1) polynucleotides may be used to produce the secreted (1) or to the nucleic acids may be used to produce the secreted (1) by inserting protein. (1) proteins and polynucleotides may be used to express the diagnose and treat immune/haematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AAK87403 to AAK87694 represent human immune/haematopoietic antigen genomic sequences from the present invention. AAK54992 to AAK5950 and AAM82169 represent sequences used in the exemplification of the present invention.

Query Match	79.0%	Score 15.8	DB 22	Length 749
Best Local Similarity	89.5%	Pred. No. 2.7e+02		
Matches	17	Conservative	0	Mismatches 2
				Indels 0
				Gaps 0
1	GGACAGCGCTGACACGCC	19		
558	GGACATGCGCCGACACAGCC	576		

RESULT 19  
AAK74243

ID AAK74243 standard; DNA; 749 BP.  
XX AAK74243;  
AC  
XX  
DT 07-NOV-2001 (first entry)  
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29055.  
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
OS Homo sapiens.  
XX  
XX MO200157182-A2.  
PD  
XX 09-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190706.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 11-JUL-2000; 2000US-0216880.  
PR 14-JUL-2000; 2000US-0217487.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226679.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 20-OCT-2000; 2000US-024617.  
PR 01-NOV-2000; 2000US-0246175.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249219.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 17-NOV-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.

PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251855.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.  
 PA

PI Rosen CA, Barash SC, Ruben SM;  
 XX

DR WPI; 2001-483426/52.  
 XX

PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and  
 PT metastasis -  
 XX

PS Disclosure; SEQ ID NO 29055; 3071pp + Sequence listing; English.  
 XX

CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patient's own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting  
 CC the nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent the  
 CC diagnose and treat immune/hematopoietic-related diseases, especially  
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/hematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
 CC represent sequences used in the exemplification of the present invention.  
 XX

SQ Sequence 749 BP; 201 A; 185 C; 137 G; 226 T; 0 other;

Query Match 79.0%; Score 15.8; DB 22; Length 749;  
 Best Local Similarity 89.5%; Pred. No. 2.7e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCC 19

Db 558 GGACATGCCCGACAGCC 576

Search completed: August 1, 2003, 13:55:14  
 Job time: 7.69868 secs



Sun Aug 3 09:03:43 2003

US-09-834-291-10.rnpb

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:05:17 ; Search time 5.18607 Seconds  
(without alignments)  
7955.924 Million cell updates/sec

US-09-834-291-10

Title: 1 ggcacagcctgacacgcca 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*

3: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*

4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*

5: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*

6: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*

7: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*

8: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*

9: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*

10: /cgn2\_6/ptodata/1/pubpna/US09C\_NEW\_PUB.seq:\*

11: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*

12: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*

13: /cgn2\_6/ptodata/1/pubpna/US10C\_NEW\_PUB.seq:\*

14: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*

15: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	US-09-834-291-5	Sequence 5, Appli
2	20	100.0	20	US-09-834-291-10	Sequence 10, Appli
3	20	100.0	20	US-09-834-291-24	Sequence 24, Appli
4	20	100.0	266	US-09-834-291-32	Sequence 32, Appli
5	20	100.0	720	US-09-834-291-3	Sequence 3, Appli
6	20	100.0	2380	US-09-834-291-4	Sequence 4, Appli
7	20	100.0	2827	US-09-834-291-1	Sequence 1, Appli
8	20	100.0	3212	US-09-834-291-6	Sequence 6, Appli
9	18.4	92.0	20	US-09-834-291-7	Sequence 7, Appli
10	16.8	84.0	20	US-09-834-291-9	Sequence 9, Appli
11	16.8	84.0	20	US-10-027-632-174763	Sequence 9, Appli
12	16.8	84.0	2940917	US-09-728-444-1101	Sequence 1101, Ap
13	16.4	82.0	516	US-09-983-965-2743	Sequence 2743, Ap
14	16.4	80.0	433	US-09-983-965-2743	Sequence 2743, Ap
15	15.8	79.0	165	US-09-878-574-9518	Sequence 9518, Ap
16	15.8	79.0	487	US-10-027-632-87503	Sequence 87503, A

17	15.8	79.0	487	13	US-10-027-632-316681	Sequence 316681,
18	15.8	79.0	88421	10	US-09-976-059-1	Sequence 1, Appli
19	15.8	79.0	155074	10	US-10-026-188-6	Sequence 6, Appli
20	15.8	79.0	3309400	13	US-09-738-762-1	Sequence 1, Appli
21	15.8	79.0	9025608	14	US-10-156-761-1	Sequence 1, Appli
22	15.4	77.0	272	10	US-09-983-965-3975	Sequence 3975, Ap
23	15.4	77.0	401	9	US-09-795-668-493	Sequence 493, App
24	15.4	77.0	401	9	US-09-946-807-493	Sequence 493, App
25	15.4	77.0	401	13	US-10-027-632-263366	Sequence 263366
26	15.4	77.0	673	10	US-09-764-877-966	Sequence 966, App
27	15.4	77.0	634	10	US-09-823-8304-557	Sequence 557, App
28	15.4	77.0	845	10	US-09-764-877-3925	Sequence 3925, Ap
29	15.4	77.0	860	10	US-09-764-877-3925	Sequence 3925, Ap
30	15.4	77.0	860	10	US-09-764-877-3925	Sequence 3925, Ap
31	15.4	77.0	872	13	US-10-027-632-127734	Sequence 127734,
32	15.4	77.0	872	13	US-10-027-632-127734	Sequence 127734,
33	15.4	77.0	2155	13	US-10-027-632-102862	Sequence 102862,
34	15.4	77.0	2155	13	US-10-027-632-102862	Sequence 102862,
35	15.4	77.0	3038	14	US-10-128-714-258	Sequence 258, App
36	15.4	77.0	3604	14	US-10-128-714-5258	Sequence 5258, App
37	15.4	77.0	34668	11	US-09-900-449A-3	Sequence 449A, Ap
38	15.4	77.0	1503841	9	US-09-795-686-1	Sequence 795, Ap
39	15.4	77.0	1503841	9	US-09-795-686-1	Sequence 795, Ap
40	15.4	77.0	1503841	10	US-09-946-807-1	Sequence 1, Appli
41	15.2	76.0	424	10	US-09-983-965-1862	Sequence 1862, Ap
42	15.2	76.0	428	10	US-09-983-965-2134	Sequence 2134, Ap
43	15.2	76.0	502	10	US-09-783-590-2161	Sequence 2161, Ap
44	15.2	76.0	663	13	US-10-027-632-187538	Sequence 187538,
45	15.2	76.0	663	13	US-10-027-632-187539	Sequence 187539,

ALIGNMENTS

RESULT 1

US-09-834-291-5

Sequence 5, Application US/09834291

Patent No. US20020042064A1

GENERAL INFORMATION:

APPLICANT: Krammer, Peter

APPLICANT: Muller-Schilling, Martina

APPLICANT: Oren, Moshe

TITLE OF INVENTION: P53 Binding Areas

FILE REFERENCE: 4121-122

CURRENT APPLICATION NUMBER: US/09/834, 291

PRIOR FILING DATE: 2001-08-21

PRIOR FILING DATE: 1999-10-18

PRIOR APPLICATION NUMBER: PCT/DE99/03343

PRIOR FILING DATE: 1998-10-16

NUMBER OF SEQ ID NOS: 32

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 5

LENGTH: 20

TYPE: DNA

ORGANISM: Homo Sapiens

US-09-834-291-5

Query Match 100.0%; Score 20; DB 9; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.2;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACAGCCTGACACGCA 20

DB 1 GGCACAGCCTGACACGCA 20

RESULT 2

US-09-834-291-10

Sequence 10, Application US/09834291

Patent No. US20020042064A1

GENERAL INFORMATION:

APPLICANT: Krammer, Peter

APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 10  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-10

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAGCCCTGACAAGCCA 20  
DB 1 GGACAAGCCCTGACAAGCCA 20

RESULT 3  
US-09-834-291-24  
Sequence 24, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 24  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-24

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAGCCCTGACAAGCCA 20  
DB 1 GGACAAGCCCTGACAAGCCA 20

RESULT 4  
US-09-834-291-32  
Sequence 32, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343

PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 32  
LENGTH: 266  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-32

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 266;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAGCCCTGACAAGCCA 20  
DB 160 GGACAAGCCCTGACAAGCCA 179

RESULT 5  
US-09-834-291-2  
Sequence 2, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 2  
LENGTH: 720  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-2

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 720;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAGCCCTGACAAGCCA 20  
DB 160 GGACAAGCCCTGACAAGCCA 179

RESULT 6  
US-09-834-291-3  
Sequence 3, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 3  
LENGTH: 2380

Sun Aug 3 09:03:43 2003

us-09-834-291-10.rmpb

Page 3

TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-3

Query Match 100.0%; Score 20; DB 9; Length 2380;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20  
DB 1820 GGACAAAGCCCTGACAAAGCCA 1839

RESULT 7  
US-09-834-291-4  
Sequence 4, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2827  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-4

Query Match 100.0%; Score 20; DB 9; Length 2827;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20  
DB 2267 GGACAAAGCCCTGACAAAGCCA 2286

RESULT 8  
US-09-834-291-1  
Sequence 1, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 1  
LENGTH: 3212  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-1

Query Match 100.0%; Score 20; DB 9; Length 3212;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20  
DB 2659 GGACAAAGCCCTGACAAAGCCA 2678

RESULT 9  
US-09-834-291-6  
Sequence 6, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 6  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-6

Query Match 92.0%; Score 18.4; DB 9; Length 20;  
Best Local Similarity 95.0%; Pred. No. 7.4;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20  
DB 1 GGAAAGCCCTGACAAAGCCA 20

RESULT 10  
US-09-834-291-7  
Sequence 7, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 7  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-7

Query Match 84.0%; Score 16.8; DB 9; Length 20;  
Best Local Similarity 90.0%; Pred. No. 46;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20  
DB 1 GGAAAGCCCTGACAAAGCCA 20

RESULT 11

US-09-834-291-9  
Sequence 9, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 9  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-9

Query Match 84.0%; Score 16.8; DB 9; Length 20;  
Best Local Similarity 90.0%; Pred. No. 46;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GGACAGCCCTGACAGCCA 20  
DB 1 GGACAGCCCTGACAGCCA 20

RESULT 12  
US-10-027-632-174763/C  
Sequence 174763, Application US/10027632  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
PRIOR FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 174763  
LENGTH: 2940917  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1).....(2940917)  
OTHER INFORMATION: n = A,T,C or G  
US-10-027-632-174763

Query Match 84.0%; Score 16.8; DB 13; Length 2940917;  
Best Local Similarity 90.0%; Pred. No. 46;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GGACAGCCCTGACAGCCA 20

DB 2746843 GGACAGCCCTGAATGCCA 2746824

RESULT 13  
US-09-728-444-1101/C  
Sequence 1101, Application US/09728444  
Patent No. US20020161207A1  
GENERAL INFORMATION:  
APPLICANT: Friedlich, Glenn  
APPLICANT: Zambrowicz, Brian  
TITLE OF INVENTION: No. US20020161207A1 Marine Polynucleotide Sequences  
FILE REFERENCE: Lex-0100-USA  
CURRENT APPLICATION NUMBER: US/09/728,444  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: US 60/168,360  
PRIOR FILING DATE: 1999-12-01  
NUMBER OF SEQ ID NOS: 1206  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1101  
LENGTH: 516  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)....(516)  
OTHER INFORMATION: n = A,T,C or G  
US-09-728-444-1101

Query Match 82.0%; Score 16.4; DB 10; Length 516;  
Best Local Similarity 94.4%; Pred. No. 74;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GACAGCCCTGACAGCC 19  
DB 100 GACAGCCCTGACAGCC 83

RESULT 14  
US-09-983-965-2743/C  
Sequence 2743, Application US/09983965  
Patent No. US20020137160A1  
GENERAL INFORMATION:  
APPLICANT: Warren, Wesley C.  
APPLICANT: Tao, Ningling  
APPLICANT: Byatt, John C.  
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
FILE REFERENCE: 37-21(10297)C  
CURRENT APPLICATION NUMBER: US/09/983,965  
PRIOR FILING DATE: 2001-10-26  
PRIOR APPLICATION NUMBER: US 09/465,231  
PRIOR FILING DATE: 1999-12-15  
PRIOR APPLICATION NUMBER: US 60/113,678  
PRIOR FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 5912  
SEQ ID NO 2743  
LENGTH: 433  
TYPE: DNA  
ORGANISM: Bos taurus  
FEATURE:  
OTHER INFORMATION: Clone ID: 13-LIB3058-025-Q1-K1-D1  
US-09-983-965-2743

Query Match 80.0%; Score 16; DB 10; Length 433;  
Best Local Similarity 100.0%; Pred. No. 12e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 CAGCCCTGACAGCC 19

Sun Aug 3 09:03:43 2003

us-09-834-291-10.rnpb

Page 5

Db 102 CAAGCCCTGACAAGCC 87

RESULT 15

US-09-878-574-9518  
; Sequence 9518, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 9518  
; LENGTH: 165  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: 701102580H1  
US-09-878-574-9518

Query Match 79.0%; Score 15.8; DB 10;  
Best Local Similarity 89.5%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GGACAAGCCCTGACAAGCC 19  
DB 131 GGACAAGCTCGACAAGCC 149

Search completed: August 1, 2003, 13:36:56  
Job time: 17.1861 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:37:27 ; Search time 78.1573 Seconds  
(without alignments)  
10468.541 Million cell updates/sec

Title: US-09-834-291-12

Perfect score: 20  
Sequence: 1 agagatgcccaactgttct 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: GenEmbl:  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_ov:\*  
5: gb\_ov:\*  
6: gb\_ov:\*  
7: gb\_ov:\*  
8: gb\_ov:\*  
9: gb\_ov:\*  
10: gb\_ov:\*  
11: gb\_ov:\*  
12: gb\_ov:\*  
13: gb\_ov:\*  
14: gb\_ov:\*  
15: gb\_ov:\*  
16: gb\_ov:\*  
17: gb\_ov:\*  
18: gb\_ov:\*  
19: gb\_ov:\*  
20: gb\_ov:\*  
21: gb\_ov:\*  
22: gb\_ov:\*  
23: gb\_ov:\*  
24: gb\_ov:\*  
25: gb\_ov:\*  
26: gb\_ov:\*  
27: gb\_ov:\*  
28: gb\_ov:\*  
29: gb\_ov:\*  
30: gb\_ov:\*  
31: gb\_ov:\*  
32: gb\_ov:\*  
33: gb\_ov:\*  
34: gb\_ov:\*  
35: gb\_ov:\*  
36: gb\_ov:\*  
37: gb\_ov:\*  
38: gb\_ov:\*  
39: gb\_ov:\*  
40: gb\_ov:\*  
41: gb\_ov:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	20	100.0	20	6	AX026100	AX026100 Sequence
2	20	100.0	40	6	AX026106	AX026106 Sequence
3	20	100.0	40	6	AX026114	AX026114 Sequence
4	20	100.0	2344	6	HSCD955FR	X87625 H. sapiens C
5	20	100.0	2827	6	AX026092	AX026092 Sequence
6	20	100.0	3212	6	AX026089	AX026089 Sequence
7	20	100.0	45121	6	AX695635	AX695635 Sequence
8	20	100.0	187313	9	AL157394	AL157394 Human DNA
9	19	95.0	193566	2	AC142173	AC142173 Human DNA
10	18.4	92.0	74951	2	AL513364	AL513364 Human DNA
11	18.4	92.0	163453	9	AC099561	AC099561 Homo sapi
12	18.4	92.0	165071	2	BX005378	BX005378 Danto rer
13	18.4	92.0	168522	2	AC024302	AC024302 Homo sapi
14	18.4	92.0	182635	2	BX323040	BX323040 Homo sapi
15	18.4	92.0	236685	2	AC084744	AC084744 Mus muscu
16	18	90.0	110000	2	AC099203	AC099203 Rattus no
17	18	90.0	215492	2	AC128141	AC128141 Rattus no
18	17.4	87.0	645	8	AT290110	AT290110 Zea mays
19	17.4	87.0	782	8	AT290110	AT290110 Zea mays
20	17.4	87.0	16157	1	AE001670	AE001670 Chlamydia
21	17.4	87.0	17395	1	AE002255	AE002255 Chlamydia
22	17.4	87.0	20563	2	AC015202	AC015202 Chlamydia
23	17.4	87.0	110000	6	AR310754	AR310754 Drosophila
24	17.4	87.0	11251	5	AL591593	AL591593 Drosophila
25	17.4	87.0	129606	2	AC018486	AC018486 Drosophila
26	17.4	87.0	147727	9	AC079160	AC079160 Homo sapi
27	17.4	87.0	160121	2	BX088586	BX088586 Danto rer
28	17.4	87.0	163448	10	AC116579	AC116579 Mus muscu
29	17.4	87.0	170729	2	AC118688	AC118688 Mus muscu
30	17.4	87.0	175326	2	AL954772	AL954772 Danto rer
31	17.4	87.0	181053	3	AC022346	AC022346 Drosophila
32	17.4	87.0	200098	5	AL935317	AL935317 Zebrafish
33	17.4	87.0	212154	2	BX323453	BX323453 Danto rer
34	17.4	87.0	222313	2	AC094871	AC094871 Rattus no
35	17.4	87.0	235483	2	AC134161	AC134161 Rattus no
36	17.4	87.0	254946	2	AL844587	AL844587 Danto rer
37	17.4	87.0	258988	2	AC123558	AC123558 Mus muscu
38	17.4	87.0	259868	2	AC103322	AC103322 Rattus no
39	17.4	87.0	299868	1	AE003496	AE003496 Drosophila
40	17.4	87.0	325865	1	AP002548	AP002548 Chlamydia
41	17	85.0	34200	9	AB022785	AB022785 Homo sapi
42	17	85.0	37068	8	AC067938	AC067938 Neurospor
43	17	85.0	43556	8	AC067937	AC067937 Neurospor
44	17	85.0	74043	2	AC121551	AC121551 Mus muscu
45	17	85.0	96376	2	AC007626	AC007626 Homo sapi

## ALIGNMENTS

RESULT 1  
LOCUS AX026100 20 bp DNA linear PAT 16-SEP-2000  
DEFINITION Sequence 12 from Patent DE19847779.  
ACCESSION AX026100  
VERSION AX026100.1 GI:10187531  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Euteleostomi; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE  
AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.  
TITLE Novel receptor dna useful for identifying apoptosis-modulating  
substances potentially useful for cancer chemotherapy

## JOURNAL

Patent: DE 19847779-C 12 03-FEB-2000;  
DEUTSCHES KREBSFORSCH (DE)

## FEATURES

Location/Qualifiers  
1. .20

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 6 a 4 c 4 g 6 t  
ORIGIN

## Query Match

Best Local Similarity 100.0%; Score 20; DB 6; Length 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20  
1 AGAGATGCCCAACTGTTT 20  
Db 1 AGAGATGCCCAACTGTTT 20

## RESULT 2

AX026106

LOCUS AX026106 40 bp DNA linear PAT 16-SEP-2000  
DEFINITION Sequence 18 from Patent DE19847779.  
ACCESSION AX026106  
VERSION AX026106.1 GI:10187537  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)

## REFERENCE

AUTHORS

TITLE

1 Mueller-Schilling, M., Kramer, P. and Oren, M.  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

## JOURNAL

Novel receptor dna useful for identifying apoptosis-modulating  
substances potentially useful for cancer chemotherapy  
Patent: DE 19847779-C 18 03-FEB-2000;  
DEUTSCHES KREBSFORSCH (DE)

## FEATURES

Location/Qualifiers  
1. .40

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 14 a 11 c 5 g 10 t  
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 40;  
Best Local Similarity 100.0%; Pred. No. 1.7;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20  
11 AGAGATGCCCAACTGTTT 30  
Db 11 AGAGATGCCCAACTGTTT 30

## RESULT 3

AX026114

LOCUS AX026114 40 bp DNA linear PAT 16-SEP-2000  
DEFINITION Sequence 26 from Patent DE19847779.  
ACCESSION AX026114  
VERSION AX026114.1 GI:10187545  
KEYWORDS  
SOURCE Homo sapiens (human)

## REFERENCE

AUTHORS

TITLE

1 Mueller-Schilling, M., Kramer, P. and Oren, M.  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

## JOURNAL

Novel receptor dna useful for identifying apoptosis-modulating  
substances potentially useful for cancer chemotherapy  
Patent: DE 19847779-C 26 03-FEB-2000;  
DEUTSCHES KREBSFORSCH (DE)

## FEATURES

Location/Qualifiers  
1. .40

BASE COUNT 14 a 11 c 5 g 10 t  
ORIGIN

BASE COUNT 14 a 11 c 5 g 10 t  
ORIGIN

## Query Match

Best Local Similarity 100.0%; Score 20; DB 6; Length 40;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20  
11 AGAGATGCCCAACTGTTT 30  
Db 11 AGAGATGCCCAACTGTTT 30

## RESULT 4

HSCD955FR

LOCUS HSCD955FR 2344 bp DNA linear PRI 05-FEB-1997  
DEFINITION H.sapiens CD95 gene 5' flanking region.  
ACCESSION X87625  
VERSION X87625.1 GI:902311  
KEYWORDS beta interferon; CD95 gene; silencer.  
SOURCE Homo sapiens (human)

## REFERENCE

AUTHORS

TITLE

1 Rudert, F., Vasser, E., Forbes, L., Lindridge, E., Wang, Y. and  
Watson, J.  
Identification of a silencer, enhancer, and basal promoter region  
in the human CD95 (Fas/APO-1) gene  
DNA Cell Biol. 14 (11), 931-937 (1995)  
96069539  
7576179  
2 (bases 1 to 2344)  
Rudert, F.H.  
Direct Submission  
Submitted (26-MAY-1995) F.H. Rudert, Genesis Research &  
Development, Corporation Ltd., PO Box 50, Auckland, NEW ZEALAND  
Overlaps with X81335, & X82279-X82286.  
Location/Qualifiers  
1. .2344

## JOURNAL

MEBLINE

PUBMED 96069539  
7576179  
2 (bases 1 to 2344)  
Rudert, F.H.  
Direct Submission  
Submitted (26-MAY-1995) F.H. Rudert, Genesis Research &  
Development, Corporation Ltd., PO Box 50, Auckland, NEW ZEALAND  
Overlaps with X81335, & X82279-X82286.  
Location/Qualifiers  
1. .2344

## REFERENCE

AUTHORS

1 Rudert, F., Vasser, E., Forbes, L., Lindridge, E., Wang, Y. and  
Watson, J.  
Identification of a silencer, enhancer, and basal promoter region  
in the human CD95 (Fas/APO-1) gene  
DNA Cell Biol. 14 (11), 931-937 (1995)  
96069539  
7576179  
2 (bases 1 to 2344)  
Rudert, F.H.  
Direct Submission  
Submitted (26-MAY-1995) F.H. Rudert, Genesis Research &  
Development, Corporation Ltd., PO Box 50, Auckland, NEW ZEALAND  
Overlaps with X81335, & X82279-X82286.  
Location/Qualifiers  
1. .2344

## COMMENT

FEATURES

Source

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal



ACCESSION AX026092.1 GI:10187523  
 VERSION  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 Mueller-Schilling, M., Krammer, P. and Oren, M.  
 AUTHORS Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy  
 TITLE Patent: DE 1984779-C 4 03-FEB-2000;  
 JOURNAL DEUTSCHES KREBSFORSCH (DE)  
 FEATURES Location/Qualifiers  
 Source 1..2827  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 728 a 676 c 657 g 766 t  
 ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 2827;  
 Best Local Similarity 100.0%; Pred. No. 2.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGAGATGCCCAACTGTTT 20  
 |||||  
 44 AGAGATGCCCAACTGTTT 63

RESULT 6  
 LOCUS AX026089 3212 bp DNA linear PAT 16-SEP-2000  
 DEFINITION Sequence 1 from Patent DE1984779.  
 LOCUS AX026089  
 VERSION AX026089.1 GI:10187520  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 Mueller-Schilling, M., Krammer, P. and Oren, M.  
 AUTHORS Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy  
 TITLE Patent: DE 1984779-C 1 03-FEB-2000;  
 JOURNAL DEUTSCHES KREBSFORSCH (DE)  
 FEATURES Location/Qualifiers  
 Source 1..3212  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 778 a 784 c 809 g 841 t  
 ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 3212;  
 Best Local Similarity 100.0%; Pred. No. 2.7;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGAGATGCCCAACTGTTT 20  
 |||||  
 44 AGAGATGCCCAACTGTTT 63

RESULT 7  
 LOCUS AX695635 45121 bp DNA linear PAT 31-MAR-2003  
 DEFINITION Sequence 1262 from Patent WO03008583.  
 LOCUS AX695635  
 VERSION AX695635.1 GI:29418787  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM

REFERENCE 1 Morris, D. W. and Engelhard, E. K.  
 AUTHORS Novel compositions and methods for cancer  
 TITLE Patent: WO 03008583-A 1262 30-JUN-2003;  
 JOURNAL Sagres Discovery (US)  
 FEATURES Location/Qualifiers  
 Source 1..45121  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 13226 a 8836 c 9010 g 14049 t  
 ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 45121;  
 Best Local Similarity 100.0%; Pred. No. 3.5;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGAGATGCCCAACTGTTT 20  
 |||||  
 Db 8037 AGAGATGCCCAACTGTTT 8056

RESULT 8  
 LOCUS AL157394 187313 bp DNA linear PRI 22-AUG-2001  
 DEFINITION Human DNA sequence from clone RP11-399019 on chromosome 10, complete sequence.  
 LOCUS AL157394  
 VERSION AL157394.15 GI:15384622  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 187313)  
 AUTHORS Blakey, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (22-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire, UK  
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 requests: clonerequest@sanger.ac.uk  
 On Aug 31, 2001 this sequence version replaced gi:14161146.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; information on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr10>  
 RP11-399019 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBACe3.6  
 This sequence is the entire insert of clone RP11-399019. The true left end of clone RP11-496H23 is at 166408 in this sequence. The true right end of clone RP11-30415 is at 18704 in this sequence.

Location/Qualifiers

```

source
1. 187313
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-399019"
/clone_1kb="RPC1-11.2"
100119. 100156
/notes="Sequence confirmed by AC015461 sequenced by WIBR."
misc_feature
100157. 100198
/notes="Sequence confirmed by WIBR."
misc_feature
105808. 105972
/notes="Sequence confirmed by WIBR."
misc_feature
105973. 105989
/notes="Sequence confirmed by WIBR."
misc_feature
105973. 105989
/notes="Sequence confirmed by WIBR."
BASE COUNT 55669 a 36398 c 36888 g 58358 t
ORIGIN
Query Match 100.0%; Score 20; DB 9; Length 187313;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGATGCCCAACTGTTT 20
Db 142012 AGAGATGCCCAACTGTTT 142031

```

```

RESULT 9
AC142173 193566 bp DNA 1linear HTG 28-MAR-2003
LOCUS Rattus norvegicus clone CH230-135C19, WORKING DRAFT SEQUENCE, 64
DEFINITION AC142173
unordered pieces.
AC142173.2 GI:29336106
HTG; HTGS_PHASE1; HTGS_DRAFT.
Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 193566)
Muzny,D.,Maritz, Metzker,M.,Lee, Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Angiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhey,C., Burch,P., Butrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Diya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregregis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,M.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,
Hollins,B., Howells,S., Huylk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpach,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kows,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshew,L., Louised,H., Lozano,R.U., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mamoud,M., Malloy,K., Mangum,A.,
Manungu,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawliny,S., McLeod,M., McNeill,T., Meenen,E., Milosavljevic,A.,
Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K.,
Morris,S., Mundasa,M., Murphy,M., Nair,L., Nankervis,C., Neill,D.,
Newton,N., Nguyen,N., Norris,S., Nwackemele,O., Okonumu,G.,
Olatunbosun,A., Pal,S., Parks,K., Pasternak,S., Paul,H.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Poindexter,A.,
Popovic,D., Primus,E., Pu,L., Puzo,M., Quroz,J., Reichlin,E.,
Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y.,
Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojka,A.,
Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Saverly,G., Scherer,S.,
Scott,G., Shatman,S., Shen,H., Shetty,J., Shwartsbeyn,A.,
Sisson,I., Sitter,C.D., Smas,J., Sneed,A., Sodergren,E.,
Song,X.-Z., Sorelle,R., Sosa,D., Steimle,M., Strong,R., Sutton,A.,
Stavak,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S.,
Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseña,D.,
Waldron,L., Walker,B., Wang,J., Valas,R., Thomas,N., Thomas,S.,
Warren,R., Wei,X., White,F., Williams,G., Wilson,R., Wiczek,R.,
Wooden,H., Worley,K., Wright,D., Wright,R., Wu,Y., Yakub,S.,
Yen,J., Yoon,L., Yoon,Y., Yu,F., Zhang,J., Zhou,J., Zhou,X.,
Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R.,
Holt,R.A., Smith,H.O., Weinstock,G., and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 193566)
Submitted (24-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 193566)
Worley,K.C.
Direct Submission
Submitted (28-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 28, 2003 this sequence version replaced gi:29165510.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KEON
Center clone name: CH230-135C19
----- Summary Statistics
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 159659 bases at least Q40
Consensus quality: 168864 bases at least Q30
Consensus quality: 172461 bases at least Q20
Estimated insert size: 166404; sum-of-coverage estimation
Quality coverage: 3x in Q20 bases; sum-of-coverage estimation

```

```

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 64 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1025: contig of 1025 bp in length
* 1026 1125: gap of unknown length
* 1126 1257: contig of 1132 bp in length
* 1258 2357: gap of unknown length
* 2358 3598: contig of 1241 bp in length
* 3599 3699: gap of unknown length
* 3699 4877: contig of 1178 bp in length
* 4877 4977: gap of unknown length
* 4977 6088: contig of 1112 bp in length
* 6088 6189: gap of unknown length
* 6189 8079: contig of 1891 bp in length
* 8079 8080: gap of unknown length
* 8080 8180: contig of 1084 bp in length
* 8180 9264: gap of unknown length
* 9264 10952: contig of 1589 bp in length

```

10953 11052: gap of unknown length  
 11053 12469: contig of 1417 bp in length  
 12470 12569: gap of unknown length  
 12570 13999: contig of 1430 bp in length  
 14000 14099: gap of unknown length  
 14100 15332: contig of 1233 bp in length  
 15333 15432: gap of unknown length  
 15433 16637: contig of 1205 bp in length  
 16638 16737: gap of unknown length  
 16738 18048: contig of 1311 bp in length  
 18049 19368: contig of 1220 bp in length  
 19369 19468: gap of unknown length  
 19470 21333: contig of 1865 bp in length  
 21334 21433: gap of unknown length  
 21434 22762: contig of 1329 bp in length  
 22763 22862: gap of unknown length  
 22863 24894: contig of 2032 bp in length  
 24895 24994: gap of unknown length  
 24995 26516: contig of 1522 bp in length  
 26517 28758: contig of 2142 bp in length  
 28759 28858: gap of unknown length  
 28859 30914: contig of 2056 bp in length  
 30915 31014: gap of unknown length  
 31015 32359: contig of 1345 bp in length  
 32360 32459: gap of unknown length  
 32460 34080: contig of 1621 bp in length  
 34081 34180: gap of unknown length  
 34181 35184: contig of 1004 bp in length  
 35185 35284: gap of unknown length  
 35285 36895: contig of 1611 bp in length  
 36896 36995: gap of unknown length  
 36996 38112: contig of 1116 bp in length  
 38112 38211: gap of unknown length  
 38212 39375: contig of 1164 bp in length  
 39376 39475: gap of unknown length  
 39476 41610: contig of 2135 bp in length  
 41611 41710: gap of unknown length  
 41710 42750: contig of 1040 bp in length  
 42751 42850: gap of unknown length  
 42851 45186: contig of 2336 bp in length  
 45187 45286: gap of unknown length  
 45287 46671: contig of 1385 bp in length  
 46672 46771: gap of unknown length  
 46772 48826: contig of 2055 bp in length  
 48827 48926: gap of unknown length  
 48927 50979: contig of 2053 bp in length  
 50980 51079: gap of unknown length  
 51080 52880: contig of 1701 bp in length  
 52881 52980: gap of unknown length  
 52981 54986: contig of 2078 bp in length  
 54987 55085: gap of unknown length  
 55086 56859: contig of 1801 bp in length  
 56860 56959: gap of unknown length  
 56960 59033: contig of 2074 bp in length  
 59034 59133: gap of unknown length  
 59134 61078: contig of 1945 bp in length  
 61079 61178: gap of unknown length  
 61179 64211: contig of 3033 bp in length  
 64212 64311: gap of unknown length  
 64312 66389: contig of 2078 bp in length  
 66390 66489: gap of unknown length  
 66490 69371: contig of 2882 bp in length  
 69372 69471: gap of unknown length  
 69472 71543: contig of 2072 bp in length  
 71544 71643: gap of unknown length  
 71644 73921: contig of 2278 bp in length  
 73922 74021: gap of unknown length  
 74022 77070: contig of 3049 bp in length  
 77071 77170: gap of unknown length  
 77171 80276: contig of 3106 bp in length  
 80277 80376: gap of unknown length

Query Match 83531: contig of 3155 bp in length  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 AGAGATGCCCAACGTTT 19  
 DB 167412 AGAGATGCCCAACGTTT 167430

RESULT 10  
 LOCUS AL513364  
 DEFINITION Human DNA sequence from clone RP11-480N10 on chromosome 1, complete sequence.  
 ACCESSION AL513364  
 VERSION AL513364.10 GI:16973038  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Almeida, J.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-NOV-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clones@wtsanger.ac.uk  
 SOURCE humquerry@sanger.ac.uk  
 ORGANISM Homo sapiens

COMMENT  
 On Nov 16, 2001 this sequence version replaced gi:16944088.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SwissProt; Tr, TrEMBL; Mp, MOPREP; Information on the MOPREP database can be found at

http://www.sanger.ac.uk/projects/C\_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr1  
 RP11-480N10 is from the library RPI1-11.2 constructed by the group of Pieter de Jong. For further details see  
 http://www.chori.org/bacpac/home.htm

VECTER: PBACe3.6  
 IMPORTANT: This sequence is not the entire insert of clone RP11-480N10. It may be shorter because we sequence overlapping sections only once, except for a short overlap.  
 The true right end of clone RP11-480N10 is at 74951 in this sequence. The true left end of clone RP11-77G8 is at 42605 in this sequence. The true right end of clone RP11-54132 is at 2000 in this sequence.

## FEATURES

Location/Qualifiers

	Query Match	92.0%	Score 18.4;	DB 9;	Length 74951;
	Best Local Similarity	95.0%	Pred.No. 3;		
	Matches 19;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0
Qy	1 AGAGATGCCCAACTGTTT	20			
Db	51780 AGAATGCCCAACTGTTT	51799			

Genome Center  
 Center: University of Washington Genome Center  
 Center Code: UWGC  
 Web site: <http://www.genome.washington.edu>  
 Contact: uwgchtgs@u.washington.edu  
 Drafting Center: SC

---

Project Information  
 Center project name: Chr-1  
 Center clone name: RP11-193R16 (sc0121)

---

Summary Statistics  
 Sequencing vector: plasmid, L08752, 87% of reads  
 Chemistry: Dye-terminator B7, 68% of reads  
 Chemistry: Dye-terminator Big Dye, 32% of reads  
 Assembly program: Phrap, version 0.990319  
 Consensus quality: 163426 bases at least Q40

3' : RP4-722L13 (UWGC:sc0558) AL365355

**Sequence Quality Assessment:**

estimates computed by the Phrap assembly program.

quality levels above 40 are expected to have less than error in 10,000 bp.

base of base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., "Phred" quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by MultiPlexComplete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

## ECORI

## Hindi II

Bajir

SeqDerMap	FngPrInt	SeqDerMap	FngPrInt	SeqDerMap	FngPrInt
8696	8565	1949	1903	15694	15824
6	<800	6382	6533	2067	208
10511	10321	512	<800	7215	744
3291	3411	449	<800	1979	198
2717	2761	2776	2753	445	<80
1023	1057	2700	2753	1292	127
4241	4333	5132	5075	1170	116
1831	1830	12787	12698	3837	381
7818	7675	1103	1091	3564	357
932	918	1317	1257	837	906
12395	12928	5920	5812	5589	5567
2809	2761	1887	1903	58	<800
183	<800	5786	5812	8040	8112
12956	12928	5002	5075	998	906
1588	1586	8584	8683	508	<800

5958	5997	9300	9335	16398	15823
1485	1456	1942	1903	4873	4875
3790	3672	529	<800	2847	2834
2032	2037	3774	3815	4216	4112
648	<800	519	<800	2429	2499
335	<800	1939	1903	2781	2834
3249	3411	471	<800	4688	4653
5628	5579	4026	3985	3544	3572
3046	3020	2288	2347	12025	11730
2433	2442	263	<800	4093	4112
2990	3020	1478	1414	2248	2274
7657	7675	7155	7408	2317	2274
1341	1321	3557	3526	5732	5567
1825	1830	2397	2347	2481	2499
143	<800	863	894	3030	3022
3099	3244	11609	11460	253	<800
3480	3411	971	894	4136	4112
9141	9053	2669	2640	1507	1495
4104	4099	1890	1903	5520	5567
976	918	8666	8683	22304	22665
115	<800	1280	1257	295	<800
3601	3672	4288	4259	1752	1743
218	<800	4227	4259	175	<800
32	<800	325	<800	3783	3814
10290	10321	2186	2178	3843	3814
5337	5272	6371	6282	1589	1562
8075	7675	859	894		
2662	2761	2630	2640		
1733	1710	2009	2031		
3006	3020	845	894		
2726	2761	7421	7408		
		377	<800		
		1595	1590		
		1467	1414		
		851	894		
		1214	1257		

## FEATURES

source

Location/Qualifiers  
1. 163453  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="caxon:9606"  
/chromosome="1"

Query Match 92.0%; Score 18.4; DB 9; Length 163453;  
Best Local Similarity 95.0%; Pred. No. 36;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAATGCCCAACTGTTT 20  
DB 150211 AGAATGCCCAACTGTTT 150192

RESULT 12 BX005378 165071 bp DNA linear HTG 03-JAN-2003

LOCUS BX005378 Danio rerio clone DKEY-40H20, \*\*\* SEQUENCING IN PROGRESS \*\*\*

DEFINITION unordered pieces.

ACCESSION BX005378 GI:27497342

VERSION HTG: HTGS PHASE1; HTGS DRAFT.

KEYWORDS Danio rerio (zebrafish)

SOURCE Danio rerio

ORGANISM Danio rerio; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 165071)

REFERENCE McLay, K.

AUTHORS Direct Submission

TITLE Submitted (02-JAN-2003) Wellcome Trust Sanger Institute, Hinxton,

JOURNAL Cambridge, CB10 1SA, UK. E-mail enquiries: zfac@sanger.ac.uk

COMMENT Clone requests: clonerequest@sanger.ac.uk

On Jan 4, 2003 this sequence version replaced gi:27475575.

Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: zfac@sanger.ac.uk

Project Information

Center project name: zK40H20

Summary Statistics

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 161469 bases at least Q40

Consensus quality: 162178 bases at least Q30

Consensus quality: 162708 bases at least Q20

Insert size: 164371; sum-of-contigs

Insert size: 163675; 5.4% error; agarose-fp

Quality coverage: 6.43x in Q20 bases; sum-of-contigs Quality

Coverage: 6.54x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently

consists of 8 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

1 36150: contig of 36150 bp in length

1 36250: gap of 100 bp

1 36251: contig of 6192 bp in length

1 42443: gap of 100 bp

1 42543: contig of 33312 bp in length

1 42543: contig of 33312 bp in length

```

misc_feature      | /organism="Danio rerio"  
                   | /mol_type="genomic DNA"  
                   | db_xref="taxon:7955"  
                   | /clone_id="DKEY-40H20"  
                   | /clone_lib="DanicKey"  
                   | 1..36150  
/note="assembly_fragment:00533  
fragment_chain:1"  
36251..42442  
/note="assembly_fragment:00652  
fragment_chain:1"  
42543..75854  
/note="assembly_fragment:01251  
fragment_chain:1"  
75955..86826  
/note="assembly_fragment:01739  
fragment_chain:1"  
86927..131894  
/note="assembly_fragment:01951  
fragment_chain:2"  
131995..134989  
/note="assembly_fragment:00031  
fragment_chain:2"  
135090..147854  
/note="assembly_fragment:01256  
fragment_chain:2"  
147855..165071  
/note="assembly_fragment:01836  
fragment_chain:2"  

misc_count       | 54498 a  
origin           | C 29325 g 50484 t 705 others
```

```
Query Match          92.0%; Score 18.4; DB 2; Length 165071;
Best Local Similarity 95.0%; Fred. No. 36;
Matches    19; Conservative    0; Mismatches    1; Indels    0; Gaps    0.
```

[illegible]

TITLE  
JOURNAL  
COMMENT

Choepeil, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,  
Daxariello, K., Dewar, K., Diaz, J., Dodge, S., Domingo, M., Doyle, M.,  
Fenster, J., Ferreira, P., FitzHugh, W., Forrest, C., Gage, D.,  
Galagan, J., Gadyana, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heatord, A., Hottin, L.,  
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kamm, L.,  
Kleian, D., Landers, T., Laroque, K., Lehotsky, J., Levins, R.,  
Lien, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M.,  
McEwan, P., McGuirk, A., McKernan, K., McPheters, R., Meldrum, J.,  
Meneus, L., Mihov, T., Miranda, C., Mlenga, V., Morrow, J.,  
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., O'Neil, J.,  
Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C.,  
Riley, R., Rogov, P., Roehman, D., Roy, A., Santos, R., Schauer, S.,  
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Subramanian, A., Talama, J., Tasseff, S., Thodou, C., Tirrell, A.,  
Travers, M., Triggillo, J., Vassilev, H., Viet, R., Vo, A., Wilson, B.,  
Xu, X., Wyman, D., Ye, W. J., Young, G., Zaimoun, J., Zimmer, A. and  
Zody, M.

**Direct Submission**  
Submitted (28-PEB-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 14, 2000 this sequence version replaced g1:7108032.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/hw/RepeatMasker/>

----- Genome Center  
Center. Whitehead Institute  
----- [www.repeatmasker.html](http://www.repeatmasker.html)

Center code: WIBR  
Web site: <http://www.wibr.org>

Contact: [sequence\\_submissions@rockefeller.edu](mailto:sequence_submissions@rockefeller.edu)

Project Information  
Center project name: T437C

Center clone name: 23\_I\_23

Sequencing vector: M13; M7781

Assembly program: Phran: vector 3  
Dye-terminator Big Dye; 100% of reads

consensus quality: 155724 bases at least Q40  
consensus quality: 153433

Consensus quality: 164971 bases at least Q30  
Insert size: 1000 bases at least Q20

Insert size: 166822; sum-of-confidence

Quality coverage: 4.0 in Q20 bases; agarose-fp

----- This is a "work-in-progress" document, subject to change without notice.

consists of 18 contigs. The true order of the

arbitrary. Gaps between the contigs are

record will be undated with a

soon as it is available and the accession number preserved

1007: contig of 1007 bp in Tenth

2229: contig of 1133 bp  
1108 gap of 100 bp

Line	Description	Amount
2329	gap of 100 bp	
2330	containing	
4120		

4121	4220: gap of 100 bp	4221
4221	4220: gap of 100 bp	4221

6094 6193: gap of 100 bp

12756: contig of 6463 bp in length  
12756: gap of 100 bp

19311: contig of 6555 bp in length  
19411: 332,251,333,334,335,336,337,338,339,340,341,342,343,344,345,346,347,348,349,350,351,352,353,354,355,356,357,358,359,360,361,362,363,364,365,366,367,368,369,370,371,372,373,374,375,376,377,378,379,380,381,382,383,384,385,386,387,388,389,390,391,392,393,394,395,396,397,398,399,400,401,402,403,404,405,406,407,408,409,410,411,412,413,414,415,416,417,418,419,420,421,422,423,424,425,426,427,428,429,430,431,432,433,434,435,436,437,438,439,440,441,442,443,444,445,446,447,448,449,450,451,452,453,454,455,456,457,458,459,460,461,462,463,464,465,466,467,468,469,470,471,472,473,474,475,476,477,478,479,480,481,482,483,484,485,486,487,488,489,490,491,492,493,494,495,496,497,498,499,500,501,502,503,504,505,506,507,508,509,510,511,512,513,514,515,516,517,518,519,520,521,522,523,524,525,526,527,528,529,530,531,532,533,534,535,536,537,538,539,540,541,542,543,544,545,546,547,548,549,550,551,552,553,554,555,556,557,558,559,560,561,562,563,564,565,566,567,568,569,570,571,572,573,574,575,576,577,578,579,580,581,582,583,584,585,586,587,588,589,590,591,592,593,594,595,596,597,598,599,600,601,602,603,604,605,606,607,608,609,610,611,612,613,614,615,616,617,618,619,620,621,622,623,624,625,626,627,628,629,630,631,632,633,634,635,636,637,638,639,640,641,642,643,644,645,646,647,648,649,650,651,652,653,654,655,656,657,658,659,660,661,662,663,664,665,666,667,668,669,670,671,672,673,674,675,676,677,678,679,680,681,682,683,684,685,686,687,688,689,690,691,692,693,694,695,696,697,698,699,700,701,702,703,704,705,706,707,708,709,710,711,712,713,714,715,716,717,718,719,720,721,722,723,724,725,726,727,728,729,730,731,732,733,734,735,736,737,738,739,740,741,742,743,744,745,746,747,748,749,750,751,752,753,754,755,756,757,758,759,760,761,762,763,764,765,766,767,768,769,770,771,772,773,774,775,776,777,778,779,780,781,782,783,784,785,786,787,788,789,790,791,792,793,794,795,796,797,798,799,800,801,802,803,804,805,806,807,808,809,810,811,812,813,814,815,816,817,818,819,820,821,822,823,824,825,826,827,828,829,830,831,832,833,834,835,836,837,838,839,840,841,842,843,844,845,846,847,848,849,850,851,852,853,854,855,856,857,858,859,860,861,862,863,864,865,866,867,868,869,870,871,872,873,874,875,876,877,878,879,880,881,882,883,884,885,886,887,888,889,890,891,892,893,894,895,896,897,898,899,900,901,902,903,904,905,906,907,908,909,910,911,912,913,914,915,916,917,918,919,920,921,922,923,924,925,926,927,928,929,930,931,932,933,934,935,936,937,938,939,940,941,942,943,944,945,946,947,948,949,950,951,952,953,954,955,956,957,958,959,960,961,962,963,964,965,966,967,968,969,970,971,972,973,974,975,976,977,978,979,980,981,982,983,984,985,986,987,988,989,990,991,992,993,994,995,996,997,998,999,1000,1001,1002,1003,1004,1005,1006,1007,1008,1009,1010,1011,1012,1013,1014,1015,1016,1017,1018,1019,1020,1021,1022,1023,1024,1025,1026,1027,1028,1029,1030,1031,1032,1033,1034,1035,1036,1037,1038,1039,1040,1041,1042,1043,1044,1045,1046,1047,1048,1049,1050,1051,1052,1053,1054,1055,1056,1057,1058,1059,1060,1061,1062,1063,1064,1065,1066,1067,1068,1069,1070,1071,1072,1073,1074,1075,1076,1077,1078,1079,1080,1081,1082,1083,1084,1085,1086,1087,1088,1089,1090,1091,1092,1093,1094,1095,1096,1097,1098,1099,1100,1101,1102,1103,1104,1105,1106,1107,1108,1109,1110,1111,1112,1113,1114,1115,1116,1117,1118,1119,1120,1121,1122,1123,1124,1125,1126,1127,1128,1129,1130,1131,1132,1133,1134,1135,1136,1137,1138,1139,1140,1141,1142,1143,1144,1145,1146,1147,1148,1149,1150,1151,1152,1153,1154,1155,1156,1157,1158,1159,1160,1161,1162,1163,1164,1165,1166,1167,1168,1169,1170,1171,1172,1173,1174,1175,1176,1177,1178,1179,1180,1181,1182,1183,1184,1185,1186,1187,1188,1189,1190,1191,1192,1193,1194,1195,1196,1197,1198,1199,1200,1201,1202,1203,1204,1205,1206,1207,1208,1209,1210,1211,1212,1213,1214,1215,1216,1217,1218,1219,1220,1221,1222,1223,1224,1225,1226,1227,1228,1229,1230,1231,1232,1233,1234,1235,1236,1237,1238,1239,1240,1241,1242,1243,1244,1245,1246,1247,1248,1249,1250,1251,1252,1253,1254,1255,1256,1257,1258,1259,1260,1261,1262,1263,1264,1265,1266,1267,1268,1269,1270,1271,1272,1273,1274,1275,1276,1277,12

24024: config of 4613 bp in length

30616: contig of 6492 bp in Jaccot

35830: gap of 100 bp

35930: gap of 100 bp

45440: gap of 100 bp



Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20  
Db 121795 AGAATGCCCAACTGTTT 121814

RESULT 15  
AC084744/c  
LOCUS

DEFINITION Mus musculus clone RP23-15A13, WORKING DRAFT SEQUENCE, 35 unordered

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

COMMENT.

JOURNAL

AUTHORS

TITLE

REFERENCE

ORGANISM

KEYWORDS

VERSION

ACCESSION

DEFINITION

LOCUS

AC084744/c

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

AC084744

Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor

laboratory

Center code: CSHL

Web site: <http://www.cshl.org/genec>

Contact: [mccombie@cshl.org](mailto:mccombie@cshl.org)

Center project name: RP23-15A13

Center clone name: RP23-15A13

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 35 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1

\* 22323: contig of 22323 bp in length

\* 22432: gap of unknown length

\* 22433: contig of 19708 bp in length

\* 42141: gap of unknown length

\* 42249: contig of 15644 bp in length

\* 57893: contig of 15644 bp in length

\* 58003: gap of unknown length

\* 72098: contig of 14095 bp in length

\* 72207: gap of unknown length

\* 85432: contig of 13226 bp in length

\* 145834: gap of unknown length

\* 145941: contig of 6721 bp in length

\* 152663: gap of unknown length

\* 152770: contig of 5894 bp in length

\* 152771: gap of unknown length

\* 158665: contig of 5848 bp in length

\* 158772: gap of unknown length

\* 164621: contig of 5848 bp in length

\* 164729: gap of unknown length

\* 170567: contig of 5838 bp in length

\* 170674: gap of unknown length

\* 176255: contig of 5590 bp in length

\* 176373: gap of unknown length

\* 181546: contig of 5173 bp in length

\* 181654: gap of unknown length

\* 186815: contig of 5161 bp in length

\* 186923: gap of unknown length

\* 191172: contig of 4249 bp in length

\* 191280: gap of unknown length

\* 195508: contig of 4228 bp in length

\* 195616: gap of unknown length

\* 199209: contig of 3593 bp in length

\* 199317: gap of unknown length

\* 202692: contig of 3375 bp in length

\* 202800: gap of unknown length

\* 206146: contig of 3346 bp in length

Page 10



Sun Aug 3 09:03:53 2003

us-09-834-291-12.rpt

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:13:22 ; Search time 68.4754 Seconds  
(without alignments)  
7098.748 Million cell updates/sec

Title: US-09-834-291-12

Perfect score: 20

Sequence: 1 agagatgcccaactgtttt 20

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_estba:\*  
3: em\_estba:\*  
4: em\_estba:\*  
5: em\_estba:\*  
6: em\_estba:\*  
7: em\_estba:\*  
8: em\_estba:\*  
9: gb\_est1:\*  
10: gb\_est1:\*  
11: gb\_est1:\*  
12: gb\_est1:\*  
13: gb\_est1:\*  
14: gb\_est1:\*  
15: em\_estba:\*  
16: em\_estba:\*  
17: em\_estba:\*  
18: em\_estba:\*  
19: em\_estba:\*  
20: em\_estba:\*  
21: em\_estba:\*  
22: em\_estba:\*  
23: em\_estba:\*  
24: em\_estba:\*  
25: em\_estba:\*  
26: em\_estba:\*  
27: em\_estba:\*  
28: gb\_est1:\*  
29: gb\_est1:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	90.0	539	9	AL922459
2	17.4	87.0	282	10	BB573416
3	17.4	87.0	475	12	BM420516
4	17.4	87.0	635	9	AL963961

C	5	17.4	87.0	677	10	BG410123	BG410123	S10-8-A10
C	6	17.4	87.0	735	10	BB630820	BB630820	AG-ND-165
C	7	17.4	87.0	796	28	BH393256	BH393256	EX422214
C	8	17.2	86.0	1201	13	EX422214	EX422214	AL575246
C	9	17.2	85.0	1035	9	AL575246	AL575246	AL551725
C	10	17.2	85.0	1087	9	AL551725	AL551725	AL551725
C	11	16.8	84.0	200	28	A2766513	A2766513	1M0564H07
C	12	16.8	84.0	291	9	AA400303	AA400303	zue4a03.x
C	13	16.8	84.0	302	10	BF835617	BF835617	RC4-HT089
C	14	16.8	84.0	314	10	BF934308	BF934308	CM2-GN016
C	15	16.8	84.0	322	28	A2719811	A2719811	RPC1-24.1
C	16	16.8	84.0	326	10	BB508020	BB508020	BB508020
C	17	16.8	84.0	335	14	D60325	D60325	HUM099E06A
C	18	16.8	84.0	377	10	AM879474	AM879474	PMO-OT001
C	19	16.8	84.0	385	10	BF994425	BF994425	CM2-GN016
C	20	16.8	84.0	394	28	A2231468	A2231468	RPC1-23-6
C	21	16.8	84.0	411	28	AO173492	AO173492	HS-3193.A
C	22	16.8	84.0	466	28	BZ118730	BZ118730	CH230-460
C	23	16.8	84.0	498	10	CA332861	CA332861	hab17c10.
C	24	16.8	84.0	522	24	BZ29647	BZ29647	KD0267.R1
C	25	16.8	84.0	532	12	BM054898	BM054898	1e83907.Y
C	26	16.8	84.0	535	9	AM704890	AM704890	sk40h01.Y
C	27	16.8	84.0	546	28	AO992080	AO992080	nbe0080C
C	28	16.8	84.0	563	10	BF191086	BF191086	237803.MA
C	29	16.8	84.0	578	28	AZ909829	AZ909829	RPC1-24-2
C	30	16.8	84.0	611	29	BZ294769	BZ294769	CG1081.F1
C	31	16.8	84.0	622	28	BH762728	BH762728	EMBAC330F
C	32	16.8	84.0	634	29	CC093693	CC093693	CSU-K34.1
C	33	16.8	84.0	667	28	BZ002104	BZ002104	cee36e03.
C	34	16.8	84.0	667	29	AG037563	AG037563	pan.trog1
C	35	16.8	84.0	673	29	BZ920990	BZ920990	CH240.118
C	36	16.8	84.0	676	14	BY727273	BY727273	BY727273
C	37	16.8	84.0	684	12	BX101371	BX101371	BX101371
C	38	16.8	84.0	688	12	BJ096294	BJ096294	B096294
C	39	16.8	84.0	693	13	BH388813	BH388813	603802703
C	40	16.8	84.0	698	29	BX184556	BX184556	Danio.rer
C	41	16.8	84.0	700	28	BZ074128	BZ074128	1K984B03.
C	42	16.8	84.0	742	29	BX205069	BX205069	Danio.rer
C	43	16.8	84.0	748	28	BH988183	BH988183	ceh29a07.
C	44	16.8	84.0	783	29	AG094259	AG094259	Pan.trog1
C	45	16.8	84.0					

## ALIGNMENTS

RESULT 1  
LOCUS AL922459 539 bp mRNA linear EST 18-SRP-2002  
DEFINITION AL922459 PUR-Z1422 Danio rerio cDNA clone 111-E05-2, mRNA sequence.  
ACCESSION AL922459  
VERSION AL922459.1 GI:23189039  
KEYWORDS EST.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio  
REFERENCE Lee, S., Ruan, H., Ma, W., Xu, M., Lo, J., He, Y., Liu, F., Eun, A., Wen, Z.  
AUTHORS 1 (bases 1 to 539)  
TITLE 111-E05-2, mRNA sequence.  
JOURNAL Unpublished  
COMMENT Contact: Peng J  
15,000 unique zebrafish EST clusters from two cDNA libraries  
Lab of Functional Genomics  
Institute of Molecular and Cell Biology  
30 Medical Drive, Singapore, 117609, Singapore  
Email: pengj@imcb.a-star.edu.sg  
Clone requests: pengj@imcb.a-star.edu.sg  
Location/Qualifiers  
1..539  
/organism="Danio rerio"  
/mol\_type="mRNA"

/strain="local\_wildtype"  
/db\_xref="taxon:7955"  
/clone="111-B05-2"  
/issue\_type="whole embryo or fish"  
/dev\_stage="mixed stages"  
/clone\_11b="PJR-21+42"  
BASE COUNT 103 a 160 c 115 g 161 t  
ORIGIN

Query Match 90.0%; Score 18; DB 9; Length 539;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAGAGTCCCAACTGTT 18  
186 AAGAGTCCCAACTGTT 169

RESULT 2  
LOCUS BB573416/c

DEFINITION BB573416 RIKEN full-length enriched, adult male testis EST 30-NOV-2000  
musculus cDNA clone 4930501C15 5', mRNA sequence.

ACCESSION BB573416  
VERSION BB573416.1 GI:11469960  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 282)

REFERENCE 1 (bases 1 to 282)  
Alizawa, K., Akahira, S., Akimura, T., Arai, A., Arakawa, T., Carninci, P., Hanagaki, T., Hayatsu, N., Hirooka, T., Hirozane, T., Hodojima, Y., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kawai, J., Kojima, Y., Kono, Nomura, K., Numaizaki, R., Okazaki, Y., Miyazaki, A., Nakamura, M., Nishi, K., Sakai, K., Sasaki, D., Sato, K., Shibata, K., Owa, C., Sakai, C., Shitaki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Tanaka, T., Toya, A., Watanabe, A., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshiki, A., Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Alizawa, K. et al. 2000)  
Unpublished

COMMENT  
TITLE Laboratory for Genome Exploration Research Group, RIKEN Genomic  
JOURNAL Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suenho-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Thermolabile and thermolabile activation of thermolabile enzymes by  
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)

Tomaru, Y., Carninci, P., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Y. and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
19-44 (1999)  
Please visit our web site (http://genome.riken.go.jp) for  
further details.  
Location/Qualifiers  
1..282  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CS7BL/6J"  
/db\_xref="taxon:10090"  
/clone="4930501C15"

FEATURES  
source

/sex="male"  
/issue\_type="testis"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_11b="RIKEN full-length enriched, adult male testis  
(DH10B)"  
/note="Site 1: Sali; Site 2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5'  
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTCTTTT 3'] cDNA was  
prepared by using trehalase thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. Second strand cDNA was prepared with the  
primer adapter of sequence [5'  
GAGAGAGAGATTCGAGTTATTTAAATTAATCCCCCCCCCC 3']. cDNA  
was cloned into the XhoI and BamHI sites. Vector: a  
modified pBluescript KS(+) after bulk excision from Lambda  
FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI."

BASE COUNT 83 a 77 c 66 g 56 t  
ORIGIN

Query Match 87.0%; Score 17.4; DB 10; Length 282;  
Best Local Similarity 94.7%; Pred. No. 3e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GAGATGCCCAACTGTTT 20  
149 GAGATGCCCAACTGTTT 131

RESULT 3  
LOCUS BM420516

DEFINITION BM420516 475 bp mRNA linear EST 28-JAN-2002  
U013C07 Oryza sativa mature leaf library induced by M.grisea Oryza  
sativa cDNA clone U013C07, mRNA sequence.

ACCESSION BM420516  
VERSION BM420516.1 GI:18387318  
KEYWORDS EST.  
SOURCE Oryza sativa  
ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eriacoidae; Oryzae; Oryza.  
1 (bases 1 to 475)

REFERENCE 1 (bases 1 to 475)  
Dong, H. T., Li, D. B., Zhang, X. F., Dai, C. G., Sun, L. X., Pei, Y. X., Wu, H. F., Jiang, Y. X., Yu, P. C., Gao, Q. K. and Lou, Y. C.  
A Gene Expression Screen in Oryza sativa  
Unpublished

COMMENT  
TITLE Laboratory of Functional Genetics  
JOURNAL Bio-technology Institute of Zhejiang University  
Kaixuan Road 266#, Hangzhou, Zhejiang, P.R.China  
Tel: 0086-571-86892051  
Fax: 0086-571-86891525  
Email: htdong@zjuem.zju.edu.cn  
Seg primer: M13 forward primer.  
Location/Qualifiers  
1..475  
/organism="Oryza sativa"  
/mol\_type="mRNA"  
/db\_xref="taxon:4530"  
/clone="U013C07"  
/issue\_type="leaf"  
/dev\_stage="mature stage"  
/clone\_11b="Oryza sativa mature leaf library induced by  
M.grisea"  
/note="Vector: pSport2"

FEATURES  
source

BASE COUNT 140 a 88 c 109 g 137 t 1 others  
ORIGIN

Sciences Center (SSC), Tokai Chemical Research Laboratory,  
The Institute of Physical and Chemical Research, Kanagawa 230-0045, Japan  
1-7-22 Shuhiro-cho, Tsurumi-ku,  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@ssc.riken.go.jp/  
URL: <http://genome.gsc.riken.go.jp/>



BP 191 91006 EVRY cedex - France  
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 5162.f For  
 more information about this cluster, see :  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0DM02AE10QPl&cluster=5162.f. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Paraday Avenue Genoscope sequence ID : CS0DM02AE10QPl.

# FEATURES

source  
 1. 1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DM02Y119"  
 /issue\_type="FETAL LIVER"  
 /dev\_stage="fetal"  
 /clone\_lib="Homo sapiens FETAL LIVER"  
 /note="Organ: liver; Vector: PCWVSPT 6; 1st strand cDNA  
 was primed with a NotI-oligo (dt) primer. Five prime end  
 enriched, double-strand cDNA was digested with Not I and  
 cloned into the Not I and EcoRV sites of the PCWVSPT 6  
 vector. Library was not normalized."  
 306 a 331 c 289 g 198 t 77 others

BASE COUNT 306 a 331 c 289 g 198 t 77 others  
 ORIGIN  
 Query Match 86.0%; Score 17.2; DB 13; Length 1201;  
 Best Local Similarity 85.0%; Pred. No. 8.2e+02;  
 Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 1 AGAGATGCCCAACTGTTT 20  
 |||||  
 1128 AGGATBCCCAAYGTGTTT 1109

RESULT 9  
 AL575246 1035 bp mRNA linear EST 01-JUN-2003  
 LOCUS AL575246 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CS0DM062YB12 5-PRIME, mRNA sequence.  
 ACCESSION AL575246  
 VERSION AL575246.2 GI:31313554  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1035)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization.  
 Unpublished  
 On Feb 16, 2001 this sequence version replaced gi:12936226.  
 COMMENT  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Paraday Avenue Genoscope sequence ID : CS0DM062DA06QPl.

# FEATURES

source  
 1. 1035  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DM062YB12"  
 /issue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dt)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and EcoRV sites of the PCWVSPT 6 vector. Library was normalized."  
 sites of the PCWVSPT 6 vector. Library was normalized."

BASE COUNT 224 a 261 c 166 g 248 t 136 others  
 ORIGIN

Query Match 85.0%; Score 17; DB 9; Length 1035;  
 Best Local Similarity 100.0%; Pred. No. 9.6e+02;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 4 GATGCCCAACTGTTT 20  
 |||||  
 201 GATGCCCAACTGTTT 217

RESULT 10  
 AL551725 1087 bp mRNA linear EST 31-MAY-2003  
 LOCUS AL551725 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CS0DM062Y118 3-PRIME, mRNA sequence.  
 ACCESSION AL551725  
 VERSION AL551725.2 GI:31273541  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1087)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization.  
 Unpublished  
 On Feb 15, 2001 this sequence version replaced gi:12889952.  
 COMMENT  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Paraday Avenue Genoscope sequence ID : CS0DM062DF09NPl.

# FEATURES

source  
 1. 1087  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DM062Y118"  
 /issue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dt)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and EcoRV sites of the PCWVSPT 6 vector. Library was normalized."  
 sites of the PCWVSPT 6 vector. Library was normalized."

BASE COUNT 265 a 279 c 243 g 276 t 24 others  
 ORIGIN  
 Query Match 85.0%; Score 17; DB 9; Length 1087;  
 Best Local Similarity 100.0%; Pred. No. 9.8e+02;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 4 GATGCCCAACTGTTT 20  
 |||||  
 628 GATGCCCAACTGTTT 644

RESULT 11  
 A2766513/c 200 bp DNA linear GSS 16-FEB-2001  
 LOCUS A2766513 1M0564H07F Mouse 10kb plasmid UNGCM library Mus musculus genomic  
 DEFINITION clone UNGCM10564H07 F, genomic survey sequence.  
 ACCESSION A2766513  
 VERSION A2766513.1 GI:12883665  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 (bases 1 to 200)

Dunn, D., Aoyagi, A., Barber, M., Beacoin, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingy, A., von Niederhausen, A. and Wright, D., Weis, R.

## TITLE

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

## JOURNAL

Unpublished

## COMMENT

Contact: Robert B. Weis  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: dunn@genetics.utah.edu  
Insert Length: 1000 Std Error: 0.00  
Plate: 0564 row: H column: 07  
Seq primer: CCGTGTAAACGACGCGCAGT  
Class: plasmid ends  
High quality sequence stop: 200.

## FEATURES

source

Location/Qualifiers

1..200

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/clone="UTGCM0564H07"

/sex="male"

/lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"

/clone\_1lb="Mouse 10kb plasmid UGCM library"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## BASE COUNT

63 a 32 c 35 g 70 t

## Query Match

Best Local Similarity 84.0%; Score 16.8; DB 28; Length 200;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

## CY

1 AGAGATGCCCAACTGTTT 20

## DB

137 AGAATGCCCAACTGTTT 118

## RESULT 12

AA400303/c

LOCUS

DEFINITION

AA400303

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## REFERENCE

1 (bases 1 to 291)

Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gieseler, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Maita, M., Martin, J., Moore, B., Schellenberg, K., Seftone, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R., and Wilson, R.

## TITLE

WashU-Merck EST Project 1997

## JOURNAL

Unpublished

## COMMENT

Contact: Wilson RK  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through INM; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -28m3 rev2 87 from Amersham  
High quality sequence stop: 205.

## FEATURES

source

Location/Qualifiers

1..291

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:592684"

/db\_xref="taxon:9606"

/clone="IMAGE:742732"

/sex="male"

/lab\_host="DH10B"

/clone\_1lb="Soares testis NHT"

/note="Vector: pRT10-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATGTGAAGTGGAGCGCCGACCAATTTTCTTTTCTTTT 3'] (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT10 vector. Library constructed by Benito Soares and M. Fatima Bonaldo."

## BASE COUNT

80 a 62 c 46 g 103 t

## Query Match

Best Local Similarity 84.0%; Score 16.8; DB 9; Length 291;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

## CY

1 AGAGATGCCCAACTGTTT 20

## DB

25 AGAGCTGCCCAACTGCTT 6

## RESULT 13

BF835617

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## REFERENCE

AUTHORS

1 (bases 1 to 302)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J., and Simpson, A. J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE

PUBMED

10737800

## COMMENT

Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001

## FEATURES

## Source

Email: asimpson@ludwig.org.br  
This sequence was derived from the PAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC4&ct=RC4-NT0890-151100-017-ell&ct=2000-11-15&ct4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 17  
High quality sequence stop: 302.  
Location/Qualifiers

## BASE COUNT

76 a 54 c 65 g 107 t

## ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 302;  
Best Local Similarity 90.0%; Pred. No. 6.4e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AGAGATGCCCAACTGTTT 20  
105 AGAATGCGCAACTGTTT 124

RESULT 14  
BF994308 314 bp mRNA linear EST 23-JAN-2001

LOCUS CM2-GN0165-021100-487-803 GN0165 Homo sapiens cDNA, mRNA sequence.

DEFINITION BF994308

VERSION BF994308.1 GI:12400631

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R.,

1 (bases 1 to 314)

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bate, G.S., Simpson, D.H.,

Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.O.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE

20202663

10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the PAPESP/LICR Human Cancer Genome

## FEATURES

## Source

Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&ct=CM2-GN0165-021100-487-803&ct=2000-11-02&ct4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 8  
High quality sequence stop: 143.  
Location/Qualifiers

## BASE COUNT

71 a 75 c 74 g 94 t

## ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 314;  
Best Local Similarity 90.0%; Pred. No. 6.5e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AGAGATGCCCAACTGTTT 20  
150 AGACATGCCCAACTGTTT 169

## RESULT 15

## LOCUS

## DEFINITION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## Class: BAC ends.

## Location/Qualifiers

## 1..322

## /organism="Mus musculus"

## /mol\_type="genomic DNA"

## /strain="C57BL/6J"

## /db\_xref="taxon:10090"

## /clone="RPC1-24-161H14"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

## /sex="Male"

```
/cell_type="Spleen/Brain"  
/clone_lib="Rpci-24"  
/note="Vector: pTARBAC1; Site 1: BamH1; Site 2: BamH1;  
Rpci-24 Mouse BAC Library produced by Pieter de Jong. The  
library was cloned in the pTARBAC1 cloning vector at the  
BamH1 sites using MboI partially digested male C57BL/6J  
DNA."  
BASE COUNT 69 a 81 c 57 g 115 t  
ORIGIN
```

```
Query Match 84.0%; Score 16.8; DB 28; Length 322;  
Best Local Similarity 90.0%; Pred. No. 6.6e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AGAGATGCCCAACTGTTT 20  
||| ||||| ||||| |||||  
Db 91 AGAATGCCCAAGTGT 110
```

Search completed: August 1, 2003, 20:43:03  
Job time : 73.4754 secs



Novel receptor DNA useful for identifying compounds useful for cancer chemotherapy

XX Claim 2; Fig 5; 12pp; German.  
 XX  
 CC This invention describes a novel p53-binding region of a human CD95  
 CC receptor DNA molecule. The p53-binding region, or a vector containing  
 CC it, can be used to screen for apoptosis-modulating substances  
 CC potentially useful for cancer chemotherapy. This sequence represents a  
 CC fragment of the human CD95 receptor promoter which is capable of  
 CC binding p53.  
 CC  
 XX

SO Sequence 20 BP; 6 A; 4 C; 4 G; 6 T; 0 other;  
 Query Match  
 Best Local Similarity 100.0%; Score 20; DB 21; Length 20;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dy 1 AGAGATGCCCAAACTGTTT 20  
 1 AGAGATGCCCAAACTGTTT 20  
 Db 1 AGAGATGCCCAAACTGTTT 20

## RESULT 2

ABST77493/C  
 ID ABST77493 standard; CDNA; 732 BP.  
 AC ABST77493;  
 XX  
 DT 12-DEC-2002 (first entry)  
 XX  
 DB Frog embryonic gene sequence Q9925901.  
 XX  
 KW Frog; ss; embryonic development; developmental disorder;  
 KW microarray; cell differentiation.  
 XX  
 OS Xenopus laevis.  
 XX  
 PN US2002081610-A1.  
 PD 27-JUN-2002.  
 XX  
 PF 23-JUL-2001; 2001US-0910943.  
 XX  
 PR 21-JUL-2000; 2000US-219658P.  
 XX  
 PA (UFRQ) UNIV ROCKEFELLER.  
 XX  
 PI Hemmati-Brivanlou A, Altman CR;  
 XX WPI; 2002-626534/67.  
 DR  
 XX  
 PT  
 PT to identify genes involved in embryonic development, to identify  
 PT different types of embryonic cells, and to diagnose developmental  
 PT disorders  
 XX  
 PS Claim 1; Page 763; 823pp; English.

XX The invention relates to a nucleic acid array, where each coordinate  
 CC contains a single nucleic acid species having one of 770 nucleotide  
 CC sequences (appearing as ABST76747-ABST77516) a of a Xenopus embryonic  
 CC product, or its complement or hybridizable fragment of not less than  
 CC 20 contiguous nucleotides of one of those sequences. Also included are  
 CC detecting differential expression of embryonic genes, comprising:  
 CC (a) contacting a nucleic acid array of embryonic genes, comprising:  
 CC embryonic but not mature cells with nucleic acids from sample and  
 CC control cells; and (b) detecting differential hybridization of nucleic  
 CC acids from the sample cells relative to the control cells; and detecting  
 CC defects in development, comprising: (a) contacting nucleic acids from  
 CC products known to play a fundamental role in the development process; and  
 CC (b) detecting a difference in expression of a fundamental gene in the  
 CC sample cells relative to a standard. The invention is useful to identify  
 CC genes involved in embryonic development and related processes such as

CC cell differentiation. This would be useful for diagnosing developmental  
 CC disorders and for identifying different types of embryonic cells.  
 CC The present sequence is one of the 770 Xenopus embryonic CDNA sequences.  
 XX  
 SO Sequence 732 BP; 208 A; 154 C; 200 G; 169 T; 1 other;

Query Match  
 Best Local Similarity 87.0%; Score 17.4; DB 24; Length 732;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Dy 2 GAGATGCCCAAACTGTTT 20  
 101 GAGATGCCCAAACTGTTT 83  
 Db 101 GAGATGCCCAAACTGTTT 83

## RESULT 3

AA91990/C  
 ID AA91990 standard; DNA; 1230025 BP.  
 AC AA91990;  
 XX  
 DT 13-SEP-1999 (first entry)  
 XX  
 DB Nucleotide sequence of the complete genome of Chlamydia pneumoniae.  
 XX  
 KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;  
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
 KW vaccine; neutralising epitope; ss.  
 XX  
 OS Chlamydia pneumoniae.  
 XX  
 PN WO927105-A2.  
 PD 03-JUN-1999.  
 XX  
 PF 20-NOV-1998; 98WO-IB01890.  
 XX  
 PR 04-NOV-1998; 98US-0107078.  
 XX  
 PR 21-NOV-1997; 97FR-0014673.  
 XX  
 PA (GBST) GENSET.  
 XX  
 PI Griffiths R;  
 XX  
 DR WPI; 1999-357842/30.  
 XX  
 PT  
 PT Genome sequence of Chlamydia pneumoniae  
 XX  
 PS Claim 1; Page 291-611; 1912pp; English.

XX The present sequence represents the complete genome of Chlamydia  
 CC pneumoniae, and encodes proteins AA934584-Y35879. C. pneumoniae causes  
 CC respiratory disease such as pneumonia and bronchitis and is thought  
 CC to be a contributing factor in heart disease, sarcoidosis, sinusitis,  
 CC purulent otitis media, erythema nodosum or pharyngitis. The polypeptides  
 CC encoded by the open reading frames of the C. pneumoniae genome (see  
 CC AA934584-Y35879) can be used in immunogenic compositions as vaccines.  
 CC Vectors containing C. pneumoniae nucleotide sequences can also be  
 CC used as immunogenic compositions, especially where the vector directs  
 CC the expression of a neutralising epitope of C. pneumoniae.  
 XX  
 SO Sequence 1230025 BP; 367213 A; 249833 C; 249013 G; 363589 T; 377 other;

Query Match  
 Best Local Similarity 87.0%; Score 17.4; DB 20; Length 1230025;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Dy 2 GAGATGCCCAAACTGTTT 20  
 1034868 GAGATGCCCAAACTGTTT 1034860  
 Db 1034868 GAGATGCCCAAACTGTTT 1034860

## RESULT 4

AAS51323/c  
 ID AAS51323 standard; DNA; 234 BP.  
 XX  
 AC AAS51323;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Enterococcus faecalis DNA for cellular proliferation protein #100.  
 XX  
 DE Enterococcus faecalis DNA for cellular proliferation gene;  
 XX  
 KM Antisense; ds; prokaryotic cellular proliferation gene;  
 XX  
 KM antibiotic; antibacterial; drug design.  
 XX  
 OS Enterococcus faecalis.  
 XX  
 PN WO200170955-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 21-MAR-2001; 2001WO-US09180.  
 XX  
 PR 21-MAR-2000; 2000US-191078P.  
 XX  
 PR 23-MAY-2000; 2000US-206848P.  
 XX  
 PR 26-MAY-2000; 2000US-207727P.  
 XX  
 PR 23-OCT-2000; 2000US-242578P.  
 XX  
 PR 27-NOV-2000; 2000US-253625P.  
 XX  
 PR 22-DEC-2000; 2000US-257931P.  
 XX  
 PR 16-FEB-2001; 2001US-269308P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Haselbeck R, Ohlsen KL, Zyekind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 DR WPI; 2001-611495/70.  
 DR P-PSDB; AAU33464.  
 XX  
 PT New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids -  
 XX  
 PS Claim 27; Seq ID No 3905; 511P; English.  
 XX  
 CC The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*  
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery  
 CC programmes. The antisense nucleic acid sequence is also useful to screen  
 CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence encodes an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 234 BP; 84 A; 33 C; 46 G; 71 T; 0 other;  
 XX  
 Query Match 84.0%; Score 16.8; DB 23; Length 234;  
 Best Local Similarity 90.0%; Pred. No. 63;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 5

QY 1 AGAGATGCCCAACTGTTT 20  
 DB 87 AGAAATGCACAAACTGTTT 68

AAS53070/c  
 ID AAS53070 standard; DNA; 306 BP.  
 XX  
 AC AAS53070;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Enterococcus faecalis DNA for cellular proliferation protein #498.  
 XX  
 DE Enterococcus faecalis DNA for cellular proliferation gene;  
 XX  
 KM Antisense; ds; prokaryotic cellular proliferation gene;  
 XX  
 KM antibiotic; antibacterial; drug design.  
 XX  
 OS Enterococcus faecalis.  
 XX  
 PN WO200170955-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 21-MAR-2001; 2001WO-US09180.  
 XX  
 PR 21-MAR-2000; 2000US-191078P.  
 XX  
 PR 23-MAY-2000; 2000US-206848P.  
 XX  
 PR 26-MAY-2000; 2000US-207727P.  
 XX  
 PR 23-OCT-2000; 2000US-242578P.  
 XX  
 PR 27-NOV-2000; 2000US-253625P.  
 XX  
 PR 22-DEC-2000; 2000US-257931P.  
 XX  
 PR 16-FEB-2001; 2001US-269308P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Haselbeck R, Ohlsen KL, Zyekind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 DR WPI; 2001-611495/70.  
 DR P-PSDB; AAU35211.  
 XX  
 PT New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids -  
 XX  
 PS Claim 27; Seq ID No 6707; 511P; English.  
 XX  
 CC The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*  
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery  
 CC programmes. The antisense nucleic acid sequence is also useful to screen  
 CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence encodes an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 306 BP; 117 A; 40 C; 65 G; 84 T; 0 other;  
 XX

RESULT 6

QY 1 AGAGATGCCCAACTGTTT 20  
 DB 105 AGAAATGCACAAACTGTTT 86



KW cardiovascular disorder; respiratory disorder; neurological disease;  
 KW infection; reproductive system disorder; immunosuppressive;  
 KW antitubercular; antineoplastic; antiproliferative; cytostatic; cardiant;  
 KW vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;  
 KW virucide; fungicide; ophthalmological; gene; ss.  
 XX Homo sapiens.  
 XX MO200228877-A1.  
 XX 11-APR-2002.  
 XX 17-JAN-2001; 2001WO-US01432.  
 XX 29-SEP-2000; 2000US-236326P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;  
 XX Moore PA, Wei P, Edner R, Duan DR, Shi Y, Choi GH, Fiscella M;  
 XX Ni J;  
 XX WPI: 2002-340092/37.  
 XX P-PSDB; AAO17218.  
 XX Isolated nucleic acid molecule encoding a human secreted protein is  
 XX used in preventing, treating or ameliorating a medical condition -  
 XX Claim 1; Page 478; 538pp; English.  
 XX The present invention provides the protein and coding sequences of a  
 XX number of human secreted proteins. These can be used in the treatment of  
 XX cancer, immune diseases, haemostatic disorders, hyperproliferative  
 XX disorders, renal disorders, cardiovascular disorders, respiratory  
 XX disorders, neurological diseases, infections and reproductive system  
 XX disorders. The present sequence is a coding sequence of the invention.  
 XX Sequence 409 BP; 106 A; 62 C; 74 G; 163 T; 4 other;  
 XX  
 XX Query Match 84.0%; Score 16.8; DB 24; Length 409;  
 XX Best Local Similarity 90.0%; Pred. No. 69;  
 XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 XX  
 XX 1 AGAGATGCCCAACTGTTT 20  
 XX |||||  
 XX 228 AGAATGCGCAACTGTTT 247  
 XX  
 XX RESULT 9  
 XX AAT03501/C  
 XX ID AAT03501 standard; DNA; 661 BP.  
 XX  
 XX AAT03501;  
 XX 12-JUN-1996 (first entry)  
 XX  
 XX Papilloma virus major capsid protein DNA.  
 XX  
 XX HP-Virus 49; papilloma virus; major;  
 XX capsid protein; plasmid VS40-7; DSM 9135; diagnosis;  
 KW skin carcinomas; therapy; vaccination; ds.  
 XX  
 XX Papilloma virus.  
 XX  
 XX Key Location/Qualifiers  
 XX mat\_peptide 1..660  
 XX /\*tag= a  
 XX  
 XX DE4415743-A1.  
 XX 09-NOV-1995.  
 XX 04-MAY-1994; 94DE-4415743.

XX 04-MAY-1994; 94DE-4415743.  
 XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
 XX de Villiers zur Hausen E, Leigh I, Shannan V, zur Hausen H;  
 XX WPI: 1995-383680/50.  
 XX P-PSDB; AAR88269.  
 XX DNA encoding peptide(s) of papilloma virus major capsid protein -  
 XX useful for detecting papilloma virus in skin carcinoma  
 XX Claim 1; Fig 3; 15pp; German.  
 XX AAT03501 is 78.5% homologous to HP-Virus 49, encodes AAR88269 a  
 XX peptide of papilloma virus (PV) major capsid protein and is  
 XX expressed by the plasmid VS40-7 (DSM 9135). The DNA is useful  
 XX in diagnosis, esp. to detect PV in skin carcinomas. Major capsid  
 XX proteins encoded by a PV genome contg. the DNA are useful for  
 XX therapy and/or vaccination.  
 XX Sequence 661 BP; 195 A; 118 C; 155 G; 193 T; 0 other;  
 XX  
 XX Query Match 84.0%; Score 16.8; DB 16; Length 661;  
 XX Best Local Similarity 90.0%; Pred. No. 74;  
 XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 XX  
 XX 1 AGAGATGCCCAACTGTTT 20  
 XX |||||  
 XX 567 AGAATGCGCAACTGTTT 548  
 XX  
 XX RESULT 10  
 XX AAL45853  
 XX ID AAL45853 standard; cDNA; 1569 BP.  
 XX  
 XX AAL45853;  
 XX 08-JUL-2002 (first entry)  
 XX  
 XX Human secreted protein coding sequence SEQ ID NO: 58.  
 XX  
 KW Human; secreted protein; gene therapy; cancer; immune disease;  
 KW haemostatic disorder; hyperproliferative disorder; renal disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disease;  
 KW infection; reproductive system disorder; immunosuppressive;  
 KW antitubercular; antineoplastic; antiproliferative; cytostatic; cardiant;  
 KW vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;  
 KW virucide; fungicide; ophthalmological; gene; ss.  
 XX  
 XX Homo sapiens.  
 XX  
 XX MO200228877-A1.  
 XX 11-APR-2002.  
 XX  
 XX 17-JAN-2001; 2001WO-US01432.  
 XX 29-SEP-2000; 2000US-236326P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;  
 XX Moore PA, Wei P, Edner R, Duan DR, Shi Y, Choi GH, Fiscella M;  
 XX Ni J;  
 XX WPI: 2002-340092/37.  
 XX P-PSDB; AAO17217.  
 XX Isolated nucleic acid molecule encoding a human secreted protein is  
 XX used in preventing, treating or ameliorating a medical condition -



RESULT 13  
AAH51750/c  
ID AAH51750 standard; DNA; 3001 BP.

AAH51750;  
29-AUG-2001 (first entry)

Chromosome 13q31-q33 biallelic marker containing amplicon SEQ ID 162.

sbgl; g34665; sbg2; g35017; g35018; chromosome 13q31-q33; haplotype; biallelic marker; polymorphism; schizophrenia; bipolar disorder; ds.

Homo sapiens.

WO200058510-A2.

05-OCT-2000.

30-MAR-2000; 2000WO-IB00435.

30-MAR-1999; 99US-0126903.

30-APR-1999; 99US-0131971.

30-APR-1999; 99US-0132065.

14-JUL-1999; 99US-0143928.

27-JUL-1999; 99US-0145915.

29-JUL-1999; 99US-0146452.

28-OCT-1999; 99US-0162288.

(GIST) GENSET.

Cohen D, Blumenfeld M, Chumakov I, Bouguetelerc L, Bihain B; Essioux L;

WPI; 2000-619082/59.

Polynucleotides comprising sequences from sbgl and g35018 biallelic markers are used for genotyping and detecting schizophrenia or bipolar disorder and predisposition to these disorders -

Claim 2; Page 643-644; 737PD; English.

AAH51601 represents a human genomic nucleotide sequence comprising sbgl, g34665, sbg2, g35017 and g35018 nucleic acid sequences located on the human chromosome 13q31-q33 locus. The nucleotide sequences contain biallelic markers and polymorphisms. Sequences AAH51602 - AAH51626 and AAB62907 - AAB62915 represent cDNA human sbgl cDNA sequences and protein products. AAH51627 - AAH51631 and AAB62916 - AAB62918 represent g35018 cDNA sequences and protein products. Primers AAH51632 - AAH51639 are used to isolate sbgl cDNAs, while sbgl exons from different primates are represented by sequences AAH51642 - AAH51699. Nucleotide sequences of amplicons which comprise biallelic markers located on the chromosome 13q31-q33 locus are represented in AAH51700 - AAH51817. Biallelic markers are represented in the sequences by degenerate/undefined base codes. PCR primers AAH51818 and AAH51819 are used in the isolation of sequences of the invention. The biallelic marker containing nucleotide sequences are used to determine the identity of the nucleotide at a single or multiple a sample DNA sequence. The nucleotide sequences may be labeled and used for genotyping by determining the frequency of a biallelic marker in a population an association between a genotype and a trait, a haplotype and a trait and a phenotype and a trait can be detected. The sequences can be used to determine a predisposition to or early onset of schizophrenia or bipolar disorder or a beneficial response to or side effects related to treatment against schizophrenia or bipolar disorder.

Sequence 3001 BP; 842 A; 608 C; 584 G; 966 T; 1 other;

Query Match 84.0%; Score 16.8; DB 21; Length 3001;  
Best Local Similarity 90.0%; Pred. No. 96;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AGAGATGCCCAACTGTTT 20  
db 427 AGAGCTGCCCAACTGTTT 408

RESULT 14  
AAL03245/c  
ID AAL03245 standard; DNA; 9821 BP.

AAL03245;

21-NOV-2001 (first entry)

Human reproductive system related antigen DNA SEQ ID NO: 5933.

Human; reproductive system related antigen; reproductive system disorder; cancer; gene therapy; ds.

Homo sapiens.

WO200155320-A2.

02-AUG-2001.

17-JAN-2001; 2001WO-US01339.

31-JAN-2000; 2000US-0179065.

04-FEB-2000; 2000US-0180628.

24-FEB-2000; 2000US-0184664.

02-MAR-2000; 2000US-0186350.

16-MAR-2000; 2000US-0189874.

17-MAR-2000; 2000US-0190076.

18-APR-2000; 2000US-0198123.

19-MAY-2000; 2000US-0205515.

07-JUN-2000; 2000US-0209467.

28-JUN-2000; 2000US-0215135.

30-JUL-2000; 2000US-0216647.

07-JUL-2000; 2000US-0216880.

11-JUL-2000; 2000US-0217487.

11-JUL-2000; 2000US-0217496.

14-JUL-2000; 2000US-0220963.

26-JUL-2000; 2000US-0220964.

14-AUG-2000; 2000US-0224518.

14-AUG-2000; 2000US-0225213.

14-AUG-2000; 2000US-0225214.

14-AUG-2000; 2000US-0225267.

14-AUG-2000; 2000US-0225268.

14-AUG-2000; 2000US-0225270.

14-AUG-2000; 2000US-0225447.

14-AUG-2000; 2000US-0225757.

14-AUG-2000; 2000US-0225758.

14-AUG-2000; 2000US-0225759.

14-AUG-2000; 2000US-0226279.

18-AUG-2000; 2000US-0226281.

22-AUG-2000; 2000US-0226868.

22-AUG-2000; 2000US-0227182.

23-AUG-2000; 2000US-0227009.

30-AUG-2000; 2000US-0228924.

01-SEP-2000; 2000US-0229287.

01-SEP-2000; 2000US-0229343.

01-SEP-2000; 2000US-0229344.

05-SEP-2000; 2000US-0229345.

05-SEP-2000; 2000US-0229512.

06-SEP-2000; 2000US-0230437.

06-SEP-2000; 2000US-0230438.

08-SEP-2000; 2000US-0231242.

```

PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 02-OCT-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 20-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 08-NOV-2000; 2000US-024617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246529.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249224.
PR 17-NOV-2000; 2000US-0249245.

```

---

```

PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249266.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249268.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250161.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251031.
PR 05-DEC-2000; 2000US-0251038.
PR 05-DEC-2000; 2000US-0251039.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251859.
PR 08-DEC-2000; 2000US-0251869.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0254037.
PR 05-JAN-2001; 2001US-0259678.
PR PA (HUMA-) HUMAN GENOME SCI INC.
PR XX
PR PI Rosen CA, Barash SC, Ruben SM;
PR XX WPI; 2001-465570/50.
PR DR
PR XX
PR PT Isolated nucleic acid molecule encoding a reproductive system antigen
PR PT is used in preventing, treating or ameliorating a medical condition -
PR XX
PR PS Disclosure; SEQ ID NO 5933; 1297bp + Sequence Listing; English.
PR XX
PR CC The present invention provides the protein and coding sequences of a
PR CC number of human reproductive system related antigens. These can be used
PR CC in the prevention and treatment of reproductive system disorders,
PR CC including cancer. The present sequence is a genomic sequence encoding a
PR CC protein of the invention.
PR XX
PR SQ Sequence 9821 BP; 3056 A; 1735 C; 1689 G; 3341 T; 0 other;

```

---

```

Query Match 84.0%; Score 16.8; DB 22; Length 9821;
Best Local Similarity 90.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AGAGATGCCCAACTGCTTT 20
DB 2800 AGAGTCTCCCAACTGCTTT 2781

```

---

```

RESULT 15
AAL03246/c
ID AAL03246 standard; DNA; 9824 BP.
XX
AC AAL03246;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 5934.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
XX cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.

```







Sun Aug 3 09:03:50 2003

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:05:17 ; Search time 5.18607 Seconds  
(without alignments)  
7955.924 Million cell updates/sec

Title: US-09-834-291-12  
Perfect score: 20  
Sequence: 1 agagatgcccaactgtttt 20

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0  
1439767 seqs, 1031500376 residues

Searched: 2879534

Total number of hits satisfying chosen parameters:  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	20	100.0	20	Sequence 18, Appl
2	20	100.0	40	Sequence 18, Appl
3	20	100.0	40	Sequence 26, Appl
4	20	100.0	2827	Sequence 4, Appl
5	20	100.0	3212	Sequence 1, Appl
6	17.4	87.0	732	Sequence 13, Appl
7	16.8	84.0	20	Sequence 3905, Ap
8	16.8	84.0	306	Sequence 6707, Ap
9	16.8	84.0	388	Sequence 105156, Ap
10	16.8	84.0	650	Sequence 5933, Ap
11	16.8	84.0	9821	Sequence 5933, Ap
12	16.8	84.0	9824	Sequence 89, Appl
13	16.8	84.0	20072	Sequence 1867, Ap
14	16.8	84.0	472	Sequence 254882, Ap
15	15.8	79.0	893	
16	15.8	79.0	13	

us-09-834-291-12.rnpb

17	15.8	79.0	1195	9	US-09-759-143-878
18	15.8	79.0	1195	9	US-09-780-669-878
19	15.8	79.0	1195	9	US-09-822-827-878
20	15.8	79.0	1195	10	US-09-895-793-878
21	15.8	79.0	1195	10	US-09-895-814-878
22	15.8	79.0	1195	13	US-10-012-896-878
23	15.8	79.0	1195	14	US-10-011-095-1
24	15.8	79.0	1195	14	US-10-010-667A-1
25	15.8	79.0	1195	14	US-10-205-823-396
26	15.8	79.0	2104	9	US-09-842-552-99
27	15.8	79.0	3627	14	US-10-011-095-6
28	15.8	79.0	32172	11	US-10-010-667A-6
29	15.8	79.0	2140405	13	US-09-783-590-76212
30	15.8	79.0	368	10	US-10-027-632-1742
31	15.4	77.0	646	13	US-10-027-632-142565
32	15.4	77.0	773	13	US-10-027-632-137412
33	15.4	77.0	2690	13	US-10-027-632-111780
34	15.4	77.0	110096	10	US-09-880-107-1542
35	15.4	77.0	123526	11	US-09-910-185-11
36	15.2	76.0	136	11	US-09-747-377-316
37	15.2	76.0	136	14	US-10-105-613-316
38	15.2	76.0	327	11	US-09-803-719-945
39	15.2	76.0	353	11	US-09-803-719-945
40	15.2	76.0	369	13	US-10-027-632-19177
41	15.2	76.0	377	10	US-09-867-701-2517
42	15.2	76.0	428	13	US-10-027-632-60626
43	15.2	76.0	448	13	US-10-027-632-72196
44	15.2	76.0	455	13	US-10-027-632-188985
45	15.2	76.0			

#### ALIGNMENTS

RESULT 1  
US-09-834-291-12  
Sequence 12, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT FILING DATE: 2001-08-21  
PRIORITY FILING DATE: 1999-10-18  
PRIORITY FILING DATE: 1998-10-16  
PRIORITY FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 12  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-12

Query Match: 100.0%; Score 20; DB 9; Length 20;  
Best Local Similarity: 100.0%; Pred. No. 0.5; Indels 0;  
Matches 20; Conservative 0; Mismatches 0;

QY 1 AGAGATGCCCAACTGTTT 20  
DB 1 AGAGATGCCCAACTGTTT 20  
RESULT 2  
US-09-834-291-18  
Sequence 18, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter

Sequence 878, App  
Sequence 878, App  
Sequence 878, App  
Sequence 878, App  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 396, App  
Sequence 99, Appl  
Sequence 6, Appl  
Sequence 6161, Ap  
Sequence 76212, A  
Sequence 1742, Ap  
Sequence 142565  
Sequence 137412  
Sequence 11780  
Sequence 1542, Ap  
Sequence 11, Appl  
Sequence 316, App  
Sequence 899, App  
Sequence 945, App  
Sequence 19177, A  
Sequence 2517, Ap  
Sequence 60626, A  
Sequence 73196, A  
Sequence 188985,

APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: P53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 18  
LENGTH: 40  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-18

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 40;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20  
DB 11 AGAGATGCCCAACTGTTT 30

RESULT 3  
US-09-834-291-26  
Sequence 26, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: P53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 26  
LENGTH: 40  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-26

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 40;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20  
DB 11 AGAGATGCCCAACTGTTT 30

RESULT 4  
US-09-834-291-4  
Sequence 4, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: P53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343

PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1999-10-18  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2827  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-4

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 2827;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20  
DB 44 AGAGATGCCCAACTGTTT 63

RESULT 5  
US-09-834-291-1  
Sequence 1, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: P53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 3212  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-1

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 3212;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20  
DB 44 AGAGATGCCCAACTGTTT 63

RESULT 6  
US-09-910-943-719/c  
Sequence 719, Application US/09910943  
Patent No. US20020081610A1  
GENERAL INFORMATION:  
APPLICANT: Hemmati-Brivanlou, Ali  
TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression  
FILE REFERENCE: 7529/16148U3  
CURRENT APPLICATION NUMBER: US/09/910,943  
PRIOR FILING DATE: 2001-07-23  
NUMBER OF SEQ ID NOS: 742  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 719  
LENGTH: 732  
TYPE: DNA  
ORGANISM: Xenopus laevis  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)..(732)

OTHER INFORMATION: n may be a or g or c or t/u  
US-09-910-943-719

Query Match 87.0%; Score 17.4; DB 9; Length 732;  
Best Local Similarity 94.7%; Pred. No. 24;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GAGATGCCCAACTGTTT 20  
DB 101 GAGATGCCCAACTGTTCT 83

RESULT 7  
US-09-834-291-13

Sequence 13, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT FILING DATE: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 13  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-13

Query Match 84.0%; Score 16.8; DB 9; Length 20;  
Best Local Similarity 90.0%; Pred. No. 25;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGAGTGCCCAACTGTTT 20  
DB 1 AGAGTTCCTCAAAATGTTT 20

RESULT 8  
US-09-815-242-3905/C

Sequence 3905, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT FILING DATE: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3905  
LENGTH: 234  
TYPE: DNA  
ORGANISM: Enterococcus faecalis  
US-09-815-242-3905

Query Match 84.0%; Score 16.8; DB 9; Length 234;  
Best Local Similarity 90.0%; Pred. No. 40;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGAGTGCCCAACTGTTT 20  
DB 87 AGAATGCACCAACTGTTT 68

RESULT 9  
US-09-815-242-6707/C

Sequence 6707, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT FILING DATE: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6707  
LENGTH: 306  
TYPE: DNA  
ORGANISM: Enterococcus faecalis  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(306)  
US-09-815-242-6707

Query Match 84.0%; Score 16.8; DB 9; Length 306;  
Best Local Similarity 90.0%; Pred. No. 42;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGAGTGCCCAACTGTTT 20  
DB 105 AGAATGCACCAACTGTTT 86

RESULT 10  
US-09-974-300-1452  
; Sequence 1452, Application US/09974300  
; Patent No. US20020146721A1  
; GENERAL INFORMATION:  
; APPLICANT: Berka, Randy M.  
; APPLICANT: Clausen, Ib Groth  
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
; FILE REFERENCE: 10085-500-US  
; CURRENT APPLICATION NUMBER: US/09/974,300  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 09/680,598  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/279,526  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 8481  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1452  
; LENGTH: 368  
; TYPE: DNA  
; ORGANISM: Bacillus licheniformis  
US-09-974-300-1452

Query Match 84.0%; Score 16.8; DB 10; Length 368;  
Best Local Similarity 90.0%; Pred. No. 44;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AGAGATGCCCAAACTGTTT 20  
DB 342 AGAGATGCCCAAACTGTTT 361

RESULT 11  
US-10-027-632-205156  
; Sequence 205156, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 205156  
; LENGTH: 650  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-205156

Query Match 84.0%; Score 16.8; DB 13; Length 650;  
Best Local Similarity 90.0%; Pred. No. 49;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AGAGATGCCCAAACTGTTT 20  
DB 70 AGAGATGCCCAAACTGTTT 89

RESULT 12  
US-09-764-891-5933/c  
; Sequence 5933, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5933  
; LENGTH: 9821  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-891-5933

Query Match 84.0%; Score 16.8; DB 11; Length 9821;  
Best Local Similarity 90.0%; Pred. No. 84;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AGAGATGCCCAAACTGTTT 20  
DB 2800 AGAGATGCCCAAACTGTTT 2781

RESULT 13  
US-09-764-891-5934/c  
; Sequence 5934, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5934  
; LENGTH: 9824  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-891-5934

Query Match 84.0%; Score 16.8; DB 11; Length 9824;  
Best Local Similarity 90.0%; Pred. No. 84;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AGAGATGCCCAAACTGTTT 20  
DB 2800 AGAGATGCCCAAACTGTTT 2781

RESULT 14  
US-09-070-927A-89  
; Sequence 89, Application US/09070927A  
; Patent No. US20020120116A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Kunsch  
; APPLICANT: Patrick J. Dillon  
; APPLICANT: Steven Barash  
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 982  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA

ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/070,927A  
FILING DATE: 04-May-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/046,655  
FILING DATE: 1997-05-16  
APPLICATION NUMBER: 60/044,031  
FILING DATE: 1997-05-06  
APPLICATION NUMBER: 60/066,009  
FILING DATE: 1997-11-14  
ATTORNEY/AGENT INFORMATION:  
NAME: Kenley K. Hoover  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: PB369  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 89:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20072 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 89:  
US-09-070-927A-89

Query Match 84.0%; Score 16.8; DB 10; Length 20072;  
Best Local Similarity 90.0%; Pred. No. 96;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTGTTT 20  
DB 7383 AGAATGCACAACTGTGTTT 7402

RESULT 15  
US-09-918-995-1867  
Sequence 1867, Application US/0918995  
Publication No. US20030073623A1  
GENERAL INFORMATION:  
APPLICANT: HySeq, Inc.  
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES  
FILE REFERENCE: 20411-756  
CURRENT APPLICATION NUMBER: US/09/918,995  
CURRENT FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: US/09/235,076  
PRIOR FILING DATE: 1999-01-20  
NUMBER OF SEQ ID NOS: 38054  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1867  
LENGTH: 472  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc. feature  
LOCATION: (1)...(472)  
OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-1867

Query Match 79.0%; Score 15.8; DB 11; Length 472;  
Best Local Similarity 89.5%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GAGATGCCCAACTGTGTTT 20  
DB 7383 AGAATGCACAACTGTGTTT 7402

DB 73 GAGATGCCCAACTGTGTTT 91  
Search completed: August 1, 2003, 13:36:57  
Job time : 6.18607 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 02:21:37 ; Search time 1.65066 Seconds  
(without alignments)  
5347.959 Million cell updates/sec

Title: US-09-834-291-12

Sequence: 1 agagatgcccaactgttct 20

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/PTCUS.COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	* Query Length	DB ID	Description
1	17.4	87.0	1230025 4 US-09-198-452A-1	Sequence 1, Appli
2	16.8	84.0	661 3 US-08-578-634C-3	Sequence 3, Appli
3	16.8	84.0	661 4 US-09-430-010-3	Sequence 3, Appli
4	16.8	84.0	1001 4 US-09-539-333D-162	Sequence 162, App
5	16.8	84.0	43795 3 US-08-742-185-101	Sequence 101, App
6	15.8	79.0	1195 4 US-09-323-873A-1	Sequence 1, Appli
7	15.8	79.0	3627 4 US-09-323-873A-6	Sequence 6, Appli
8	15.4	77.0	1644 3 US-09-111-752-13	Sequence 13, Appli
9	15.4	77.0	1644 3 US-09-380-061B-15	Sequence 15, Appli
10	15.4	77.0	1644 4 US-08-487-183A-13	Sequence 5, Appli
11	15.4	77.0	1704 4 US-08-460-934-5	Sequence 5, Appli
12	15.4	77.0	1704 2 US-08-782-118-5	Sequence 8, Appli
13	15.4	77.0	1908 1 US-08-460-934-8	Sequence 8, Appli
14	15.4	77.0	1908 2 US-08-782-118-8	Sequence 1, Appli
15	15.4	77.0	1664976 4 US-08-916-421B-1	Sequence 1, Appli
16	15.4	77.0	1664976 4 US-08-916-421B-1	Sequence 2038, Ap
17	15.2	76.0	450 4 US-09-134-001C-2038	Sequence 2, Appli
18	15.2	76.0	668 3 US-08-578-634C-2	Sequence 2, Appli
19	15.2	76.0	668 3 US-09-430-010-2	Sequence 753, App
20	15.2	76.0	927 4 US-09-221-017B-753	Sequence 1283, App
21	15.2	76.0	1185 4 US-09-328-352-1283	Sequence 14, Appli
22	15.2	76.0	3970 4 US-09-499-203-14	Sequence 20, Appli
23	15.2	76.0	6204 4 US-09-499-203-20	Sequence 20, Appli
24	15.2	76.0	6792 4 US-09-374-454-20	Sequence 17, Appli
25	15.2	76.0	7387 4 US-09-499-203-17	Sequence 1, Appli
26	15.2	76.0	9321 4 US-09-499-203-1	Sequence 1, Appli

28	15.2	76.0	99500 4 US-09-798-096-10	Sequence 10, Appli
29	15.2	76.0	148567 4 US-09-801-876B-3	Sequence 3, Appli
30	15.2	76.0	319608 4 US-09-539-333D-1	Sequence 1, Appli
31	15.2	76.0	319608 4 US-09-679-409-1	Sequence 435, App
32	14.8	74.0	432 4 US-09-641-638-435	Sequence 20, Appli
33	14.8	74.0	495 3 US-09-109-204-20	Sequence 19, Appli
34	14.8	74.0	495 4 US-09-490-032-20	Sequence 19, Appli
35	14.8	74.0	605 3 US-09-109-204-19	Sequence 246, App
36	14.8	74.0	605 4 US-09-490-032-19	Sequence 240, App
37	14.8	74.0	660 4 US-08-961-527-246	Sequence 5, Appli
38	14.8	74.0	1001 4 US-09-671-317-240	Sequence 5, Appli
39	14.8	74.0	2125 3 US-09-109-204-5	Sequence 1, Appli
40	14.8	74.0	2125 4 US-09-490-032-5	Sequence 1, Appli
41	14.8	74.0	3794 4 US-09-608-821-1	Sequence 1, Appli
42	14.8	74.0	3804 4 US-08-842-382-1	Sequence 577, App
43	14.8	74.0	4712 4 US-09-221-017B-577	Sequence 28, Appli
44	14.8	74.0	26654 4 US-09-564-805-28	Sequence 1, Appli
45	14.4	72.0	832 1 US-08-473-496-1	Sequence 1, Appli

## ALIGNMENTS

RESULT 1  
US-09-198-452A-1/C  
Sequence 1, Application US/09198452A  
Patent No. 6559294  
GENERAL INFORMATION:  
APPLICANT: Grifffais, R.  
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
FILE REFERENCE: 9710-003-999  
CURRENT APPLICATION NUMBER: US/09/198, 452A  
CURRENT FILING DATE: 1998-11-24  
NUMBER OF SEQ ID NOS: 6849  
SEQ ID NO 1  
LENGTH: 1230025  
TYPE: DNA  
ORGANISM: Chlamydia pneumoniae  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1) .. (15000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (15001) .. (30000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (30001) .. (45000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (45001) .. (60000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (60001) .. (75000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (75001) .. (90000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (90001) .. (105000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (105001) .. (120000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (120001) .. (135000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (135001) .. (150000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc feature  
LOCATION: (150001) .. (165000)

```

1 OTHER INFORMATION: n=a or c or g or t
2 NAME/KEY: misc feature
3 LOCATION: (165001) .. (180000)
4 OTHER INFORMATION: n=a or c or g or t
5 NAME/KEY: misc feature
6 LOCATION: (180001) .. (195000)
7 OTHER INFORMATION: n=a or c or g or t
8 NAME/KEY: misc feature
9 LOCATION: (195001) .. (210000)
10 OTHER INFORMATION: n=a or c or g or t
11 NAME/KEY: misc feature
12 LOCATION: (210001) .. (225000)
13 OTHER INFORMATION: n=a or c or g or t
14 NAME/KEY: misc feature
15 LOCATION: (225001) .. (240000)
16 OTHER INFORMATION: n=a or c or g or t
17 NAME/KEY: misc feature
18 LOCATION: (240001) .. (255000)
19 OTHER INFORMATION: n=a or c or g or t
20 NAME/KEY: misc feature
21 LOCATION: (255001) .. (270000)
22 OTHER INFORMATION: n=a or c or g or t
23 NAME/KEY: misc feature
24 LOCATION: (270001) .. (285000)
25 OTHER INFORMATION: n=a or c or g or t
26 NAME/KEY: misc feature
27 LOCATION: (285001) .. (300000)
28 OTHER INFORMATION: n=a or c or g or t
29 NAME/KEY: misc feature
30 LOCATION: (300001) .. (315000)
31 OTHER INFORMATION: n=a or c or g or t
32 NAME/KEY: misc feature
33 LOCATION: (315001) .. (330000)
34 OTHER INFORMATION: n=a or c or g or t
35 NAME/KEY: misc feature
36 LOCATION: (330001) .. (345000)
37 OTHER INFORMATION: n=a or c or g or t
38 NAME/KEY: misc feature
39 LOCATION: (345001) .. (360000)
40 OTHER INFORMATION: n=a or c or g or t
41 NAME/KEY: misc feature
42 LOCATION: (360001) .. (375000)
43 OTHER INFORMATION: n=a or c or g or t
44 NAME/KEY: misc feature
45 LOCATION: (375001) .. (390000)
46 OTHER INFORMATION: n=a or c or g or t
47 NAME/KEY: misc feature
48 LOCATION: (390001) .. (405000)
49 OTHER INFORMATION: n=a or c or g or t
50 NAME/KEY: misc feature
51 LOCATION: (405001) .. (420000)
52 OTHER INFORMATION: n=a or c or g or t
53 NAME/KEY: misc feature
54 LOCATION: (420001) .. (435000)
55 OTHER INFORMATION: n=a or c or g or t
56 NAME/KEY: misc feature
57 LOCATION: (435001) .. (450000)
58 OTHER INFORMATION: n=a or c or g or t
59 NAME/KEY: misc feature
60 LOCATION: (450001) .. (465000)
61 OTHER INFORMATION: n=a or c or g or t
62 NAME/KEY: misc feature
63 LOCATION: (465001) .. (480000)
64 OTHER INFORMATION: n=a or c or g or t
65 NAME/KEY: misc feature
66 LOCATION: (480001) .. (495000)
67 OTHER INFORMATION: n=a or c or g or t
68 NAME/KEY: misc feature
69 LOCATION: (495001) .. (510000)
70 OTHER INFORMATION: n=a or c or g or t
71 NAME/KEY: misc feature
72 LOCATION: (510001) .. (525000)
73 OTHER INFORMATION: n=a or c or g or t

```

```

1 NAME/KEY: misc feature
2 LOCATION: (525001)..(540000)
3 OTHER INFORMATION: n=a or c or g or t
4 NAME/KEY: misc feature
5 LOCATION: (540001)..(555000)
6 OTHER INFORMATION: n=a or c or g or t
7 NAME/KEY: misc feature
8 LOCATION: (555001)..(570000)
9 OTHER INFORMATION: n=a or c or g or t
10 NAME/KEY: misc feature
11 LOCATION: (585001)..(600000)
12 OTHER INFORMATION: n=a or c or g or t
13 NAME/KEY: misc feature
14 LOCATION: (600001)..(615000)
15 OTHER INFORMATION: n=a or c or g or t
16 NAME/KEY: misc feature
17 LOCATION: (615001)..(630000)
18 OTHER INFORMATION: n=a or c or g or t
19 NAME/KEY: misc feature
20 LOCATION: (630001)..(645000)
21 OTHER INFORMATION: n=a or c or g or t
22 NAME/KEY: misc feature
23 LOCATION: (645001)..(660000)
24 OTHER INFORMATION: n=a or c or g or t
25 NAME/KEY: misc feature
26 LOCATION: (660001)..(675000)
27 OTHER INFORMATION: n=a or c or g or t
28 NAME/KEY: misc feature
29 LOCATION: (675001)..(690000)
30 OTHER INFORMATION: n=a or c or g or t
31 NAME/KEY: misc feature
32 LOCATION: (690001)..(705000)
33 OTHER INFORMATION: n=a or c or g or t
34 NAME/KEY: misc feature
35 LOCATION: (705001)..(720000)
36 OTHER INFORMATION: n=a or c or g or t
37 NAME/KEY: misc feature
38 LOCATION: (720001)..(735000)
39 OTHER INFORMATION: n=a or c or g or t
40 NAME/KEY: misc feature
41 LOCATION: (735001)..(750000)
42 OTHER INFORMATION: n=a or c or g or t
43 NAME/KEY: misc feature
44 LOCATION: (750001)..(765000)
45 OTHER INFORMATION: n=a or c or g or t
46 NAME/KEY: misc feature
47 LOCATION: (765001)..(780000)
48 OTHER INFORMATION: n=a or c or g or t
49 NAME/KEY: misc feature
50 LOCATION: (780001)..(795000)
51 OTHER INFORMATION: n=a or c or g or t
52 NAME/KEY: misc feature
53 LOCATION: (795001)..(810000)
54 OTHER INFORMATION: n=a or c or g or t
55 NAME/KEY: misc feature
56 LOCATION: (810001)..(825000)
57 OTHER INFORMATION: n=a or c or g or t
58 NAME/KEY: misc feature
59 LOCATION: (825001)..(840000)
60 OTHER INFORMATION: n=a or c or g or t
61 NAME/KEY: misc feature
62 LOCATION: (840001)..(855000)
63 OTHER INFORMATION: n=a or c or g or t
64 NAME/KEY: misc feature
65 LOCATION: (855001)..(870000)
66 OTHER INFORMATION: n=a or c or g or t
67 NAME/KEY: misc feature
68 LOCATION: (870001)..(885000)
69 OTHER INFORMATION: n=a or c or g or t
70 NAME/KEY: misc feature

```

Sun Aug 3 09:03:49 2003

us-09-834-291-12.rnt

LOCATION: (885001)..(900000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (900001)..(915000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature

Query Match 87.0%; Score 17.4; DB 4; length 1230025;  
Best Local Similarity 94.7%; Pred. No. 17;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GAGATGCCCAACTGTTT 20  
Db 1034868 GAGATGCCCAACTGTTT 1034850

RESULT 2  
US-08-578-634C-3/C  
Sequence 3, Application US/08578634C

PATENT No. 6025163  
GENERAL INFORMATION:  
APPLICANT: Vladimir Shamanin

APPLICANT: Echel Michele De Villiers  
APPLICANT: Zur Hausen

APPLICANT: Irene Leigh  
APPLICANT: Harald Zur Hausen  
TITLE OF INVENTION: DNA CODING FOR A PEPTIDE OF A PAPILLOMA

TITLE OF INVENTION: VIRUS MAIN CAPSIDE PROTEIN AND USE THEREOF  
NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas

CITY: New York  
STATE: New York  
COUNTRY: USA

ZIP: 10036  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/08/578,634C  
FILING DATE: 26-JAN-1996

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:

NAME: Abrams, Samuel B.  
REGISTRATION NUMBER: 30,605  
REFERENCE/DOCKET NUMBER: 8484-007

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 493-4935  
TELEFAX: (650) 493-5556

TELETYPE: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
LENGTH: 661 base pairs  
TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear

MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: CDS

LOCATION: 1..660  
US-08-578-634C-3

Query Match 84.0%; Score 16.8; DB 3; length 661;  
Best Local Similarity 90.0%; Pred. No. 9.2;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20

Db 567 AGAGTGCACCAAGTGT 548

RESULT 3  
US-09-430-010-3/C  
Sequence 3, Application US/09430010

PATENT No. 6395512  
GENERAL INFORMATION:  
APPLICANT: Vladimir Shamanin

APPLICANT: Echel Michele De Villiers  
APPLICANT: Zur Hausen

APPLICANT: Irene Leigh  
APPLICANT: Harald Zur Hausen  
TITLE OF INVENTION: DNA CODING FOR A PEPTIDE OF A PAPILLOMA

TITLE OF INVENTION: VIRUS MAIN CAPSIDE PROTEIN AND USE THEREOF  
NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas

CITY: New York  
STATE: New York  
COUNTRY: USA

ZIP: 10036  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/430,010

FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/578,634  
FILING DATE: 26-JAN-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Abrams, Samuel B.  
REGISTRATION NUMBER: 30,605

REFERENCE/DOCKET NUMBER: 8484-007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 493-4935

TELEFAX: (650) 493-5556  
TELETYPE: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 661 base pairs

TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: CDS

LOCATION: 1..660  
US-09-430-010-3

Query Match 84.0%; Score 16.8; DB 4; length 661;  
Best Local Similarity 90.0%; Pred. No. 9.2;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGTTT 20  
Db 567 AGAGTGCACCAAGTGT 548

RESULT 4  
US-09-539-333D-162/C  
Sequence 162, Application US/09539333D

PATENT No. 6476208  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel

APPLICANT: Blumenfeld, Marta  
APPLICANT: Chumakov, Ilya

APPLICANT: Bougueleret, Lydie  
APPLICANT: Bihain, Bernard  
TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLERIC MARKERS  
FILE REFERENCE: GENSET.047AUS  
CURRENT APPLICATION NUMBER: US/09/539,333D  
PRIORITY FILING DATE: 2000-03-30  
PRIORITY FILING DATE: 1999-03-30  
PRIORITY FILING DATE: 1999-04-30  
PRIORITY FILING DATE: 1999-04-30  
PRIORITY FILING DATE: 1999-04-30  
PRIORITY FILING DATE: 1999-07-14  
PRIORITY FILING DATE: 1999-07-27  
PRIORITY FILING DATE: 1999-07-29  
PRIORITY FILING DATE: 1999-07-29  
PRIORITY FILING DATE: 1999-07-29  
PRIORITY FILING DATE: 1999-10-28  
PRIORITY FILING DATE: 1999-10-12  
SOFTWARE: Patent .pm  
SEQ ID NO 162  
LENGTH: 3001  
TYPE: DNA  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: allele  
LOCATION: 1501  
OTHER INFORMATION: 8-94-252 : polymorphic base A or G  
FEATURE:  
NAME/KEY: misc.binding  
LOCATION: 1482..1500  
OTHER INFORMATION: 8-94-252.misl  
FEATURE:  
NAME/KEY: misc.binding  
LOCATION: 1502..1521  
OTHER INFORMATION: 8-94-252.mls2, complement  
FEATURE:  
NAME/KEY: primer.bind  
LOCATION: 1250..1267  
OTHER INFORMATION: upstream amplification primer  
FEATURE:  
NAME/KEY: primer.bind  
LOCATION: 1651..1669  
OTHER INFORMATION: downstream amplification primer, complement  
FEATURE:  
NAME/KEY: misc.binding  
LOCATION: 1489..1513  
OTHER INFORMATION: 8-94-252 probe  
US-09-539-333D-162

Query Match  
Best Local Similarity 84.0%; Score 16.8; DB 4; Length 3001;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 AGAGATGCCCAAACTGTTT 20  
427 AGAGCTGCCCAAACTGCTTT 408

RESULT 5  
US-08-742-185-101  
Sequence 101 Application US/08742185  
Patent No 6020476  
GENERAL INFORMATION:  
APPLICANT: Page, David C.  
APPLICANT: Reijo, Renee

APPLICANT: Saxena, Richa  
APPLICANT: Hawkins, Trevor  
APPLICANT: Reeve, Mary Pat  
TITLE OF INVENTION: DAZ: A GENE FAMILY ASSOCIATED WITH AZOOSPERMIA  
NUMBER OF SEQUENCES: 102  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millitia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: US  
ZIP: 02173

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/742,185  
FILING DATE: 30-OCT-1996  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/690,734  
FILING DATE: 31-JUL-1996  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/310,429  
FILING DATE: 22-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: WHI94-07A2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ ID NO: 101:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43795 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-742-185-101

Query Match  
Best Local Similarity 84.0%; Score 16.8; DB 3; Length 43795;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 AGAGATGCCCAAACTGTTT 20  
22516 AGAGATTGCCCAAACTGTTT 22535

RESULT 6  
US-09-323-873A-1/c  
Sequence 1 Application US/09323873A  
Patent No. 6329503  
GENERAL INFORMATION:  
APPLICANT: Daniel E. Afar  
APPLICANT: Rene S. Hubert  
APPLICANT: Arthur B. Raitano  
APPLICANT: Douglas C. Saffran  
APPLICANT: Steve Chappell Mitchell  
TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS  
FILE REFERENCE: 129.16US02  
CURRENT APPLICATION NUMBER: US/09/323,873A  
PRIORITY FILING DATE: 1999-06-01  
PRIORITY FILING DATE: 1998-06-01  
PRIORITY FILING DATE: 1998-06-01  
PRIORITY FILING DATE: 1998-06-30  
NUMBER OF SEQ ID NOS: 32

SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 1195  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-323-873A-1

Query Match  
Best Local Similarity 79.0%; Score 15.8; DB 4; Length 1195;  
Pred. No. 35;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GAGATGCCCAACTGTTT 20  
DB 566 GAGATGCCCAACTGCTTT 548

RESULT 7  
US-09-323-873A-6/c  
Sequence 6, Application US/09323873A  
Patent No. 6329503  
GENERAL INFORMATION:  
APPLICANT: Daniel E. Afar  
APPLICANT: Rene S. Hubert  
APPLICANT: Kahan Beong  
APPLICANT: Arthur B. Kaitano  
APPLICANT: Douglas C. Saffran  
APPLICANT: Steve Chappell Mitchell  
TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS  
TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF  
FILE REFERENCE: 129.16USU2  
CURRENT APPLICATION NUMBER: US/09/323,873A  
CURRENT FILING DATE: 1999-06-01  
PRIOR APPLICATION NUMBER: 60/087,520  
PRIOR FILING DATE: 1998-06-01  
PRIOR APPLICATION NUMBER: 60/091,183  
PRIOR FILING DATE: 1998-06-30  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 3627  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-323-873A-6

Query Match  
Best Local Similarity 79.0%; Score 15.8; DB 4; Length 3627;  
Pred. No. 42;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GAGATGCCCAACTGTTT 20  
DB 596 GAGATGCCCAACTGCTTT 578

RESULT 8  
US-07-903-047-7/c  
Sequence 7, Application US/07903047  
Patent No. 5229285  
GENERAL INFORMATION:  
APPLICANT: Kajiyama, Naoki  
APPLICANT: Nakano, Eiichi  
TITLE OF INVENTION: Thermostable Luciferase Of Firefly  
TITLE OF INVENTION: Thermostable Luciferase Gene Of Firefly, No. 5229285el Recombi  
TITLE OF INVENTION: DNA, And Process For The Preparation Of Thermostable  
TITLE OF INVENTION: Luciferase Of Firefly  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/903,047  
FILING DATE: 19920623  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mierock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 7005-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212 790-9090  
TELEFAX: 212 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1644 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA to mRNA  
US-07-903-047-7

Query Match  
Best Local Similarity 77.0%; Score 15.4; DB 1; Length 1644;  
Pred. No. 60;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAGTGGCCCAACTGT 17  
DB 369 AGAGTGGCTTAACGT 353

RESULT 9  
US-09-111-752-13/c  
Sequence 13, Application US/09111752  
Patent No. 6074859  
GENERAL INFORMATION:  
APPLICANT: HIROKAWA, KOZO  
APPLICANT: KAJIYAMA, NAOKI  
APPLICANT: MURAKAMI, SEIJI  
TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND  
TITLE OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE LUMINESCENT PROTEIN  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/111,752  
FILING DATE: 08-JUL-1998  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBION, NORMAN F  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 7126-0009-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1644 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna to mRNA  
ORIGINAL SOURCE:  
ORGANISM: Luciola lateralis  
US-09-111-752-13

Query Match 77.0%; Score 15.4; DB 3; Length 1644;  
Best Local Similarity 94.1%; Pred. No. 60;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGT 17  
Db 369 AGAGATGCTTAACGT 353

RESULT 10  
US-09-380-061B-15/C

Sequence 15, Application US/09380061B  
Patent No. 6265177

GENERAL INFORMATION:  
APPLICANT: SQUIRRELL, DAVID JAMES

LOWE, PETER JOHN  
MURRAY, JAMES AUGUSTUS HENRY

TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESSES:  
ADDRESSEE: NIXON & VANDERHAYE P.C.

STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON

STATE: VIRGINIA  
COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/380,061B

FILING DATE: 25-Aug-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB98/01026

FILING DATE: 7-APR-1998

APPLICATION NUMBER: GB 9707468.8

FILING DATE: 11-APR-1997

ATTORNEY/AGENT INFORMATION:  
NAME: SADOFF, B. J.

REGISTRATION NUMBER: 36,663

REFERENCE/DOCKET NUMBER: 124-725

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000

TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1644 base pairs  
TYPE: nucleic acid

STRANDEDNESS: both  
TOPOLOGY: both

MOLECULE TYPE: cdna  
FEATURE:  
NAME/KEY: CDS

LOCATION: 1..1644  
SEQUENCE DESCRIPTION: SEQ ID NO: 15:

US-09-380-061B-15

Query Match 77.0%; Score 15.4; DB 3; Length 1644;  
Best Local Similarity 94.1%; Pred. No. 60;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGT 17

Db 369 AGAGATGCTTAACGT 353

RESULT 11  
US-08-487-183A-13/C

Sequence 13, Application US/08487183A  
Patent No. 6387675

GENERAL INFORMATION:  
APPLICANT: WOOD, Keith V.

GRUBER, Monika G.

TITLE OF INVENTION: MUTANT LUCIFERASES

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Foley & Lardner

STREET: P.O. Box 1497  
CITY: Madison

STATE: WI  
COUNTRY: USA

ZIP: 53701-1497

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,183A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/467,773

FILING DATE: 06-JUN-1995

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/177,081

FILING DATE: 03-JAN-1994

ATTORNEY/AGENT INFORMATION:  
NAME: Scanlon, William J.

REGISTRATION NUMBER: 31,136

REFERENCE/DOCKET NUMBER: 19017/166

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 258-5035

TELEFAX: (608) 258-4258

INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1644 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear

FEATURE:  
NAME/KEY: CDS

LOCATION: 1..1644

NAME/KEY: mat peptide  
LOCATION: 1..1644

US-08-487-183A-13

Query Match 77.0%; Score 15.4; DB 4; Length 1644;  
Best Local Similarity 94.1%; Pred. No. 60;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGT 17  
Db 369 AGAGATGCTTAACGT 353

RESULT 12  
US-08-460-934-5/C

Sequence 5, Application US/08460934  
Patent No. 5814465

GENERAL INFORMATION:  
APPLICANT: TATSUMI, HIROKI

FUKUDA, SATOSHI  
APPLICANT: KIKUCHI, MAMORU

APPLICANT: KOYAMA, YASUJI  
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE  
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A  
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT  
TITLE OF INVENTION: ANALYSIS METHOD  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,934  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 193798/1994  
FILING DATE: 27-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 54625/1995  
FILING DATE: 14-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 98857/1995  
FILING DATE: 24-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 7126-001-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1704 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1704  
NAME/KEY: misc feature  
LOCATION: 1..1704  
OTHER INFORMATION: /note="Nucleotide sequence of the  
OTHER INFORMATION: biotinylated firefly luciferase gene contained in recombinant  
OTHER INFORMATION: plasmid pHLf203 DNA"  
US-08-460-934-5  
Query Match 77.0%; Score 15.4; DB 1; Length 1704;  
Best Local Similarity 94.1%; Pred. No. 60;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AGAGATGCCCAACTGT 17  
Db 429 AGAGATGCCCAACTGT 413  
RESULT 13  
US-08-782-118-5/c  
; Sequence 5, Application US/08782118  
; Patent No. 5843746  
; GENERAL INFORMATION:  
; APPLICANT: TATSUMI, HIROKI  
; APPLICANT: FUKUDA, SATOSHI

APPLICANT: KIKUCHI, MAMORU  
APPLICANT: KOYAMA, YASUJI  
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE  
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A  
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT  
TITLE OF INVENTION: ANALYSIS METHOD  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/782,118  
FILING DATE: 13-JAN-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/460,934  
FILING DATE: 05-JUN-1995  
APPLICATION NUMBER: JP 193798/1994  
FILING DATE: 27-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 54625/1995  
FILING DATE: 14-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 98857/1995  
FILING DATE: 24-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 7126-001-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1704 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1704  
NAME/KEY: misc feature  
LOCATION: 1..1704  
OTHER INFORMATION: /note="Nucleotide sequence of the  
OTHER INFORMATION: biotinylated firefly luciferase gene contained in recombinant  
OTHER INFORMATION: plasmid pHLf203 DNA"  
US-08-782-118-5  
Query Match 77.0%; Score 15.4; DB 2; Length 1704;  
Best Local Similarity 94.1%; Pred. No. 60;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AGAGATGCCCAACTGT 17  
Db 429 AGAGATGCCCAACTGT 413  
RESULT 14  
US-08-460-934-8/c  
; Sequence 8, Application US/08460934  
; Patent No. 5814465

GENERAL INFORMATION:  
APPLICANT: TATSUMI, HIROKI  
APPLICANT: FUKUDA, SATOSHI  
APPLICANT: KIKUCHI, MAMORU  
APPLICANT: KOYAMA, YASUJI  
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE  
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A  
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT  
TITLE OF INVENTION: ANALYSIS METHOD  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESS: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,934  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 193798/1994  
FILING DATE: 27-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 54625/1995  
FILING DATE: 14-MAR-1995  
APPLICATION NUMBER: JP 98857/1995  
FILING DATE: 24-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 7126-001-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1908 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1..1908  
OTHER INFORMATION: /note= "The nucleotide sequence of  
OTHER INFORMATION: the biotinylated firefly luciferase gene contained in  
OTHER INFORMATION: recombinant plasmid pHLf248 DNA"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1908  
US-08-460-934-8

Query Match 77.0%; Score 15.4; DB 1; Length 1908;  
Best Local Similarity 94.1%; Pred. No. 62;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAGATGCCCAACTGT 17  
Db 369 AGAGATGCCCAACTGT 353

RESULT 15  
US-08-782-118-8/c  
; Sequence 8, Application US/08782118

Patent No. 5843746  
GENERAL INFORMATION:  
APPLICANT: TATSUMI, HIROKI  
APPLICANT: FUKUDA, SATOSHI  
APPLICANT: KIKUCHI, MAMORU  
APPLICANT: KOYAMA, YASUJI  
TITLE OF INVENTION: BIOTINYLATED FIREFLY LUCIFERASE, A GENE  
TITLE OF INVENTION: FOR BIOTINYLATED FIREFLY LUCIFERASE, A RECOMBINANT DNA, A  
TITLE OF INVENTION: PROCESS FOR PRODUCING BIOTINYLATED AND A BIOLUMINESCENT  
TITLE OF INVENTION: ANALYSIS METHOD  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESS: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/782,118  
FILING DATE: 13-JAN-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/460,934  
FILING DATE: 05-JUN-1995  
APPLICATION NUMBER: JP 193798/1994  
FILING DATE: 27-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 54625/1995  
FILING DATE: 14-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 7126-001-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1908 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1..1908  
OTHER INFORMATION: /note= "The nucleotide sequence of  
OTHER INFORMATION: the biotinylated firefly luciferase gene contained in  
OTHER INFORMATION: recombinant plasmid pHLf248 DNA"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1908  
US-08-782-118-8

Query Match 77.0%; Score 15.4; DB 2; Length 1908;  
Best Local Similarity 94.1%; Pred. No. 62;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

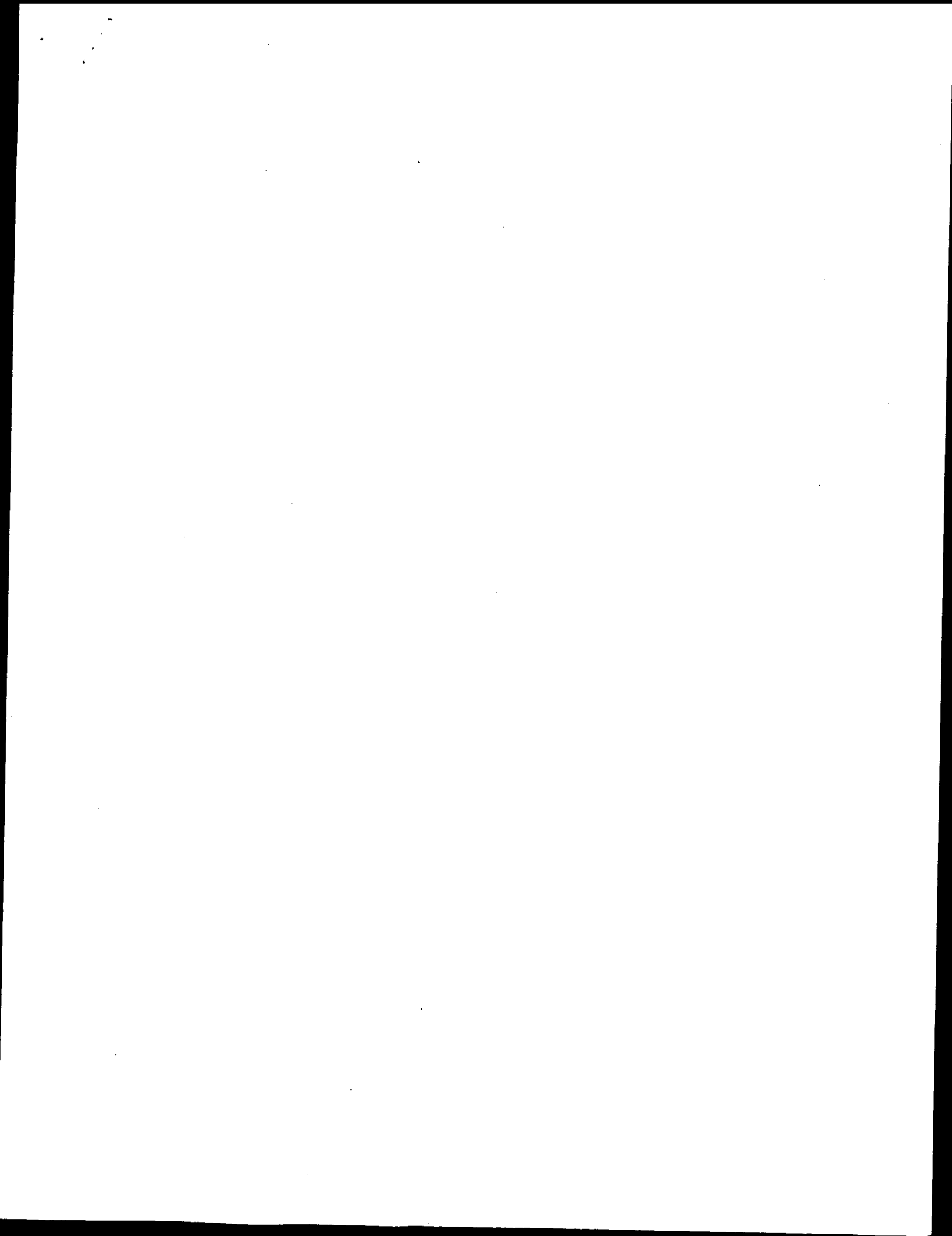
QY 1 AGAGATGCCCAACTGT 17  
Db 369 AGAGATGCCCAACTGT 353



Sun Aug 3 09:03:49 2003

Search completed: August 1, 2003, 08:37:13  
Job time : 5.65066 secs

us-09-834-291-12.rml



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:37:27 ; Search time 78.1573 Seconds  
(without alignments)  
10468.541 Million cell updates/sec

Title: US-09-834-291-14

Perfect score: 20  
Sequence: 1 aatgttcttaagcttttt 20

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2888711 segs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

GenBank1:  
1: gb Da:.\*  
2: gb Htg:.\*  
3: gb In:.\*  
4: gb Om:.\*  
5: gb Ov:.\*  
6: gb Pac:.\*  
7: gb Ph:.\*  
8: gb Pl:.\*  
9: gb Pr:.\*  
10: gb Ro:.\*  
11: gb Sts:.\*  
12: gb Sy:.\*  
13: gb Un:.\*  
14: gb Vi:.\*  
15: em Ba:.\*  
16: em Fun:.\*  
17: em Hum:.\*  
18: em In:.\*  
19: em Mu:.\*  
20: em Om:.\*  
21: em Or:.\*  
22: em Ov:.\*  
23: em Pac:.\*  
24: em Ph:.\*  
25: em Pl:.\*  
26: em Ro:.\*  
27: em Sts:.\*  
28: em Un:.\*  
29: em Vi:.\*  
30: em Htg\_hum:.\*  
31: em Htg\_inv:.\*  
32: em Htg\_other:.\*  
33: em Htg\_mus:.\*  
34: em Htg\_pln:.\*  
35: em Htg\_rod:.\*  
36: em Htg\_mam:.\*  
37: em Htg\_vrt:.\*  
38: em Sy:.\*  
39: em Htgo\_hum:.\*  
40: em Htgo\_mus:.\*  
41: em Htgo\_other:.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	* Match	Length	DB	ID	Description
1	20	100.0	20	6	AX026102	AX026102 Sequence
2	20	100.0	40	6	AX026108	AX026108 Sequence
3	20	100.0	40	6	AX026116	AX026116 Sequence
4	20	100.0	2344	6	AX026159	AX026159 Sequence
5	20	100.0	2827	6	AX026092	AX026092 Sequence
6	20	100.0	3212	6	AX026089	AX026089 Sequence
7	20	100.0	45121	6	AX026563	AX026563 Sequence
8	20	100.0	18733	9	AL157394	AL157394 Human DNA
9	19	95.0	227648	2	AC133743	AC133743 Rattus no
10	19	95.0	246752	2	AC094502	AC094502 Rattus no
11	19	95.0	264615	2	AC121457	AC121457 Rattus no
12	18.4	92.0	1454	10	RNCND1	X75207 R. norvegicu
13	18.4	92.0	1689	6	AX684916	AX684916 Sequence
14	18.4	92.0	2358	10	BC044841	BC044841 Mus muscu
15	18.4	92.0	3737	6	AX695351	AX695351 Sequence
16	18.4	92.0	3737	10	S78355	S78355 Cyl-1-cycl1
17	18.4	92.0	3751	10	RATCYCD1	D14014 Rattus rat
18	18.4	92.0	4400	10	MMU67925	U67925 Mus musculu
19	18.4	92.0	25329	8	SPAC38	Z69086 S.pombe chr
20	18.4	92.0	28486	6	AX695350	AX695350 Sequence
21	18.4	92.0	42861	2	AF105153	AF105153 Homo sapi
22	18.4	92.0	48823	2	ALB37517	ALB37517 Homo sapi
23	18.4	92.0	72414	8	AC025615	AC025615 Homo sapi
24	18.4	92.0	145290	8	AC096688	AC096688 Oryza sat
25	18.4	92.0	146516	2	AC079981	AC079981 Homo sapi
26	18.4	92.0	153040	2	AC116618	AC116618 Homo sapi
27	18.4	92.0	154480	10	AF384675	AF384675 Mus muscu
28	18.4	92.0	154613	10	AC129213	AC129213 Mus muscu
29	18.4	92.0	163102	9	AC093749	AC093749 Homo sapi
30	18.4	92.0	170131	2	AC084379	AC084379 Homo sapi
31	18.4	92.0	173690	2	AC024055	AC024055 Homo sapi
32	18.4	92.0	177312	9	AC079801	AC079801 Homo sapi
33	18.4	92.0	178130	2	AC119975	AC119975 Mus muscu
34	18.4	92.0	179198	9	AL592183	AL592183 Human DNA
35	18.4	92.0	180141	2	AC009706	AC009706 Homo sapi
36	18.4	92.0	181997	2	AC055849	AC055849 Homo sapi
37	18.4	92.0	183417	2	AC068822	AC068822 Homo sapi
38	18.4	92.0	191190	10	AC122906	AC122906 Mus muscu
39	18.4	92.0	198073	2	AC083946	AC083946 Mus muscu
40	18.4	92.0	220190	2	AC108852	AC108852 Mus muscu
41	18.4	92.0	248187	2	AC097677	AC097677 Rattus no
42	18.4	92.0	250008	2	AC095937	AC095937 Rattus no
43	18.4	92.0	260336	2	AC094614	AC094614 Rattus no
44	18	90.0	178001	10	AL772227	AL772227 Mouse DNA
45	18	90.0	185694	2	AC079505	AC079505 Mus muscu

# ALIGNMENTS

RESULT 1  
AX026102  
LOCUS AX026102  
DEFINITION Sequence 14 from Patent DE19847779.  
ACCESSION AX026102  
VERSION AX026102.1 GI:10187533  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.  
TITLE Novel receptor dna useful for identifying apoptosis-modulating  
substances potentially useful for cancer chemotherapy

Pred. No. is the number of results predicted by chance to have a

## JOURNAL

Patent: DE 19847779-C 14 03-FEB-2000;  
DEUTSCHES KREBSFORSCH (DE)

## FEATURES

Location/Qualifiers  
1. 20  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

## BASE COUNT

4 a 2 c 3 g 11 t

## Query Match

Best Local Similarity 100.0%; Score 20; DB 6; Length 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 1 AATGTTGCTTAAGCTTTT 20  
1 AATGTTGCTTAAGCTTTT 20

## RESULT 2

AX026108

LOCUS

Sequence 20 from Patent DE19847779. 40 bp DNA linear PAT 16-SEP-2000

ACCESSION

AX026108

VERSION

AX026108.1 GI:10187539

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

REFERENCE

AUTHORS

Muehler-Schilling, M., Kramer, P. and Oren, M.

JOURNAL

DEUTSCHES KREBSFORSCH (DE)

FEATURES

Location/Qualifiers

1. 40

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT

7 a 8 c 6 g 19 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 20; DB 6; Length 40;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 20

11 AATGTTGCTTAAGCTTTT 30

## RESULT 3

AX026116

LOCUS

Sequence 28 from Patent DE19847779. 40 bp DNA linear PAT 16-SEP-2000

ACCESSION

AX026116

VERSION

AX026116.1 GI:10187547

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

REFERENCE

AUTHORS

Muehler-Schilling, M., Kramer, P. and Oren, M.

JOURNAL

DEUTSCHES KREBSFORSCH (DE)

FEATURES

Location/Qualifiers

1. 40

/organism="Homo sapiens"

## BASE COUNT

7 a 8 c 6 g 19 t

## Query Match

Best Local Similarity 100.0%; Score 20; DB 6; Length 40;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 1 AATGTTGCTTAAGCTTTT 20  
11 AATGTTGCTTAAGCTTTT 30

## RESULT 4

HSCD955FR

LOCUS

H.sapiens CD95 gene 5' flanking region. 2344 bp DNA linear PRI 05-FEB-1997

ACCESSION

X87625

VERSION

X87625.1 GI:902311

KEYWORDS

beta interferon; CD95 gene; silencer.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

Rudert, F., Vasser, E., Forbes, L., Lindridge, E., Wang, Y. and Watson, J.

JOURNAL

Identification of a silencer, enhancer, and basal promoter region in the human CD95 (Fas/Apo-1) gene

DNA Cell Biol. 14 (11), 931-937 (1995)

PUBMED

96069539

REFERENCE

2 (bases 1 to 2344)

AUTHORS

Rudert, F.H.

JOURNAL

Submitted (26-MAY-1995) F.H. Rudert, Genesis Research &amp; Development, Corporation Ltd., PO Box 50, Auckland, NEW ZEALAND

TITLE

Development, Corporation Ltd., PO Box 50, Auckland, NEW ZEALAND

COMMENT

Overlaps with X81335, &amp; X82279-X82286.

FEATURES

Location/Qualifiers

1. 2344

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="10"

/map="q24.1"

/tissue\_type="placenta"

/germline

564.1337

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal

## BASE COUNT

637 a 546 c 511 g 650 t

## Query Match

Best Local Similarity 100.0%; Score 20; DB 9; Length 2344;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 1 AATGTTGCTTAAGCTTTT 20  
11 AATGTTGCTTAAGCTTTT 30

## RESULT 5

AX026092

LOCUS

Sequence 4 from Patent DE19847779. 2827 bp DNA linear PAT 16-SEP-2000

ACCESSION

AX026092

VERSION

AX026092.1 GI:10187539

KEYWORDS

beta interferon; CD95 gene; silencer.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

Rudert, F., Vasser, E., Forbes, L., Lindridge, E., Wang, Y. and Watson, J.

JOURNAL

Identification of a silencer, enhancer, and basal promoter region in the human CD95 (Fas/Apo-1) gene

DNA Cell Biol. 14 (11), 931-937 (1995)

PUBMED

96069539

REFERENCE

2 (bases 1 to 2344)

AUTHORS

Rudert, F.H.

JOURNAL

Submitted (26-MAY-1995) F.H. Rudert, Genesis Research &amp; Development, Corporation Ltd., PO Box 50, Auckland, NEW ZEALAND

TITLE

Development, Corporation Ltd., PO Box 50, Auckland, NEW ZEALAND

COMMENT

Overlaps with X81335, &amp; X82279-X82286.

FEATURES

Location/Qualifiers

1. 2344

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="10"

/map="q24.1"

/tissue\_type="placenta"

/germline

564.1337

misc\_signal

misc\_signal

misc\_signal

misc\_signal

misc\_signal

ACCESSION AX026092 GI:10187523  
 VERSION  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
 AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.  
 TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy  
 JOURNAL Patent: DE 1984779-C 4 03-FEB-2000;  
 DEUTSCHES KREBSFORSCH (DE)  
 Location/Qualifiers

FEATURES  
 source 1..2827  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 728 a 676 c 657 g 766 t  
 ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 2827;  
 Best Local Similarity 100.0%; Pred. No. 68;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 20  
 437 AATGTTGCTTAAGCTTTT 456

RESULT 6  
 AX026089 3212 bp DNA linear PAT 16-SEP-2000  
 LOCUS  
 DEFINITION Sequence 1 from Patent DE19847779.  
 ACCESSION AX026089  
 VERSION AX026089.1 GI:10187520  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
 AUTHORS Mueller-Schilling, M., Krammer, P. and Oren, M.  
 TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy  
 JOURNAL Patent: DE 1984779-C 1 03-FEB-2000;  
 DEUTSCHES KREBSFORSCH (DE)  
 Location/Qualifiers

FEATURES  
 source 1..3212  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 778 a 784 c 809 g 841 t  
 ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 3212;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 20  
 437 AATGTTGCTTAAGCTTTT 456

RESULT 7  
 AX695635 45121 bp DNA linear PAT 31-MAR-2003  
 LOCUS  
 DEFINITION Sequence 1262 from Patent WO03008563.  
 ACCESSION AX695635  
 VERSION AX695635.1 GI:29418787  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1  
 AUTHORS Morris, D.W. and Engelhard, E.K.  
 TITLE Novel compositions and methods for cancer  
 JOURNAL Patent: WO 03008583-A 1262 30-JAN-2003;  
 Segres Discovery (US)  
 Location/Qualifiers

FEATURES  
 source 1..45121  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 13226 a 8836 c 9010 g 14049 t  
 ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 45121;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 20  
 8430 AATGTTGCTTAAGCTTTT 8449

RESULT 8  
 AL157394 187313 bp DNA linear PRI 22-AUG-2001  
 LOCUS  
 DEFINITION Human DNA sequence from clone RP11-399019 on chromosome 10,  
 complete sequence.  
 ACCESSION AL157394  
 VERSION AL157394.15 GI:15384622  
 KEYWORDS  
 SOURCE HTG.  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
 AUTHORS Blakey, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (22-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 requests: clonerequest@sanger.ac.uk  
 On Aug 31, 2001 this sequence version replaced GI:1461146.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as follows unless otherwise noted: all  
 this sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em.: EMBL; Sw.:  
 SWISSPROT; Tr.: TrEMBL; Wp.: WormPeP; Information on the WormPeP  
 database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 10, constructed by the Sanger Centre Chromosome 10  
 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr10  
 RP11-399019 is from the library RPc1-11.2 constructed by the group  
 of Pieter de Jong. For further details see  
 http://www.chori.org/bacpac/home.htm  
 VECTOR: DBAC3.6  
 This sequence is the entire insert of clone RP11-399019 The true  
 left end of clone RP11-496H23 is at 166408 in this sequence. The  
 true right end of clone RP11-30415 is at 18704 in this sequence.  
 Location/Qualifiers

[illegible]

RESULT 9	AC133743/c	AC133743	22648 bp	DNA	linear	HTG 20-NOV-2002
LOCUS		Rattus norvegicus clone CH230-182C23.	***	SEQUENCING IN PROGRESS		
DEFINITION		***, 9 unordered pieces.				
ACCESSION		AC133743				
VERSION		AC133743.2	GI:25139560			
KEYWORDS		HTG; HTGS; PHASE1; HTGS; DRAFT; HTGS_ENRICHED.				
SOURCE		Rattus norvegicus (Norway rat)				
ORGANISM		Rattus norvegicus				

## REFERENCE AUTHORS

Arreola, Allen, H., Alsbrooks, S., Amn, A., Anguliano, D.,  
Aryalbecchi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaranaitake, D., Barber, M., Barnstead, M., Benahmed, F.,  
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhey, C., Burch, P., Butrell, K., Calderon, K.,  
Cardenas, V., Carter, K., Cavazos, I., Ceaear, H., Center, A.,  
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z.,  
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A.,  
Davila, M. L., Davis, C., Davy-Carroll, L., De Ande, C.,  
Delgado, O., Denison, S., Deramo, C., Ding, X., Dinh, H.,  
Diaper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B.,  
Eagan, A., Escotto, M., Eugene, C., Evans, C., Falls, T.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M.,  
Frazer, C. M., Gabriel, A., Gante, R., Garcia, A., Garner, T.,  
Gebregeorgis, E., Geer, K., Giller, R., Grady, M., Guerra, M.,  
Gunnarane, S., Haaland, W., Hamill, C., Hamilton, C.,  
Guerrata, W., Haaland, W., Hamill, C., Hamilton, C.,  
Harvey, Y., Haykal, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A.,  
Hollins, B., Howells, S., Hui, S., Hume, J., Idelbradt, D.,  
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R.,  
Karpaty, S., Kelly, S., Kelly, S., Khan, Z., King, L.,  
Kovacs, C., Kreft, C. L., Ledow, H., Levan, J., Lewis, L.,  
Li, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorenzen, H., L., Louissege, H., Lozano, R. J., Lu, X.,  
Maheshwari, M., Mahindaratne, M., Mahmoud, M.,  
Mangum, B., Mapua, P., Martin, K., Martin, R.,  
Mawhinney, S., McLeod, M. P., McNell, T. Z.,  
Mielosavljevic, A., Miner, G., Minja, E., Montemayor, J.,  
Morgan, M., Morris, K., Morris, S., Mundaas, M.,  
Nankervis, C., Neal, D., Newton, N., Nguyen, N.,  
Nkwoke, O., Okwuonu, G., Olapumusaagon, A.,  
Pal, S., Parks, K.,

REFERENCE AUTHORS TITLE JOURNAL	JOURNAL REFERENCE AUTHORS TITLE JOURNAL
1. H. J. ... 2. ... 3. ... 4. ... 5. ... 6. ... 7. ... 8. ... 9. ... 10. ... 11. ... 12. ... 13. ... 14. ... 15. ... 16. ... 17. ... 18. ... 19. ... 20. ... 21. ... 22. ... 23. ... 24. ... 25. ... 26. ... 27. ... 28. ... 29. ... 30. ... 31. ... 32. ... 33. ... 34. ... 35. ... 36. ... 37. ... 38. ... 39. ... 40. ... 41. ... 42. ... 43. ... 44. ... 45. ... 46. ... 47. ... 48. ... 49. ... 50. ... 51. ... 52. ... 53. ... 54. ... 55. ... 56. ... 57. ... 58. ... 59. ... 60. ... 61. ... 62. ... 63. ... 64. ... 65. ... 66. ... 67. ... 68. ... 69. ... 70. ... 71. ... 72. ... 73. ... 74. ... 75. ... 76. ... 77. ... 78. ... 79. ... 80. ... 81. ... 82. ... 83. ... 84. ... 85. ... 86. ... 87. ... 88. ... 89. ... 90. ... 91. ... 92. ... 93. ... 94. ... 95. ... 96. ... 97. ... 98. ... 99. ... 100. ...	1. H. J. ... 2. ... 3. ... 4. ... 5. ... 6. ... 7. ... 8. ... 9. ... 10. ... 11. ... 12. ... 13. ... 14. ... 15. ... 16. ... 17. ... 18. ... 19. ... 20. ... 21. ... 22. ... 23. ... 24. ... 25. ... 26. ... 27. ... 28. ... 29. ... 30. ... 31. ... 32. ... 33. ... 34. ... 35. ... 36. ... 37. ... 38. ... 39. ... 40. ... 41. ... 42. ... 43. ... 44. ... 45. ... 46. ... 47. ... 48. ... 49. ... 50. ... 51. ... 52. ... 53. ... 54. ... 55. ... 56. ... 57. ... 58. ... 59. ... 60. ... 61. ... 62. ... 63. ... 64. ... 65. ... 66. ... 67. ... 68. ... 69. ... 70. ... 71. ... 72. ... 73. ... 74. ... 75. ... 76. ... 77. ... 78. ... 79. ... 80. ... 81. ... 82. ... 83. ... 84. ... 85. ... 86. ... 87. ... 88. ... 89. ... 90. ... 91. ... 92. ... 93. ... 94. ... 95. ... 96. ... 97. ... 98. ... 99. ... 100. ...

COMMENT

Pasternak, S., Paul, H., Perez, A., Perez, A., Perez, L., Pfankuch, C.,  
 Plöpper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,  
 Plazio, M., Quirroz, O., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,  
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
 Rivers, C., Rodkey, T., Rojas, A., Rose, M., Roese, R., Ruiz, S. J.,  
 Sanders, W., Savary, G., Scheerer, S., Scott, G., Shatman, S., Shen, H.,  
 Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smajs, D.,  
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
 Steidle, M., Strong, R., Sutton, C., Sytek, A., Taber, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemari, K.,  
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,  
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
 Williams, G., Willison, R., Wleczyk, R., Wooden, J., Wright, D.,  
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
 Yu, E., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,  
 Weinstock, G. and Gibbs, R. A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 227648)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (18-SEP-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 227648)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (20-NOV-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Nov 20, 2002 this sequence version replaced gi:23096694.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: KAPR  
 Center clone name: CH230-182C23  
 ----- Summary Statistics  
 Assembly program: Phrap, version 0.990329  
 Consensus quality: 189225 bases at least Q40  
 Consensus quality: 193444 bases at least Q30  
 Consensus quality: 195848 bases at least Q20  
 Estimated insert size: 190572; sum-of-contigs estimation  
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_difft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_difft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 consists of 9 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1 50224: contig of 50224 bp in length  
 50225 50324: gap of unknown length  
 50325 53089: contig of 2765 bp in length

FEATURES

Source

1. 227648  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-182C23"  
129007.130011  
/note="wgs contig"  
165940.167729  
/note="wgs contig"  
198830.199863  
/note="wgs contig"  
203077.204817  
/note="wgs contig"  
59974 a 38543 c 39231 g 59843 t 30057 others

BASE COUNT 59974 a 38543 c 39231 g 59843 t 30057 others

ORIGIN

Query Match 95.0%; Score 19; DB 2; Length 227648;  
Best Local Similarity 100.0%; Pied No. 1.1e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AATGCTTAAGCTTTT 19  
15664 AATGCTTAAGCTTTT 15646

RESULT 10  
AC094502 248752 bp DNA linear HTG 09-MAY-2003  
AC094502/c  
LOCUS  
DEFINITION  
3 unordered pieces.  
AC094502 GI:30466791  
AC094502.6 HTG PHASE1; HTG ENRICHED.  
HTG: HTG PHASE1 (NOTWAY rat)  
Rattus norvegicus  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS  
1 (bases 1 to 248752)  
Muzny, D., Marie, Metzger, M., Lee, Abramson, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Albrechts, S., Amin, A., Anguiano, D.,  
Anyalebechi, V., Ayaz, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaru, A., Barber, M., Barnstead, M., Benham, F.,  
Biswal, K., Blair, D., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,  
Chavez, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, J.,  
Dalla, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dim, H., Diya, K.,  
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duvall, B., Evans, K.,  
Egan, A., Escoto, M., Evans, C., Evans, C.A., Falls, T., Fan, G.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M.,  
Gunnarsson, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A.,  
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
Kapathy, S., Kelly, S., Khan, Z., King, L., Kovari, C.,  
Kovats, C., Kraft, C.B., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorenshewa, L., Lousheed, H., Lozano, R., Lu, X., Ma, J.,  
Manshwar, M., Mahdantane, M., Mahmood, M., Mallory, K., Mangum, A.,  
Mangum, B., Margue, P., Martin, K., Martin, R., Martinez, E.,  
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,  
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,  
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,  
Nwankwelu, O., Okonou, G., Olajunmoye, A., Pal, S., Parks, K.,  
Pasternak, S., Paul, H., Perez, A., Perez, A., Peris, P., Pu, L.,  
Plopper, F., Polidexter, A., Popovic, D., Primus, E., Reigh, R.,  
Puzos, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., Riggs, F.,  
Riley, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Shen, H.,  
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,  
Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Sma, J.,  
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,  
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,  
Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,  
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, A., White, F.,  
Williams, G., Willson, R., Wleczky, R., Woodson, H., Worley, K.,  
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhou, X., Zhou, Y., von  
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
Weinstock, G. and Gibbs, R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 248752)  
Worley, K.C.  
Direct Submission  
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 248752)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On May 9, 2003 this sequence version replaced gi:24818812.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence only contigs will be indicated in the feature  
table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center Project name: GATW  
Center clone name: CH230-488  
----- Summary Statistics  
Assembly program: Atlas;  
Consensus quality: 206523 bases at least Q40  
Consensus quality: 211412 bases at least Q30  
Consensus quality: 214233 bases at least Q20  
Consensus insert size: 215006; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

```
BASE COUNT      66997 a
ORIGIN
misc_feature     /organism="Rattus norvegicus"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:10116"
                  /clone="CH230-4E8"
                  1..1697
                  /note="wgs end-extension
clone_end:"Sp6"
                2836..3642
                /note="clone_boundary
clone_end:"T7
site:EcoRI
end_sequence:BH305316"
90194..91759
/note="wgs_contig"
91810..93031
/note="wgs_contig"
152801..154271
/note="wgs_contig"
184165..185672
/note="wgs_contig"
complement(237880..238728)
/note="clone_boundary
clone_end:"Sp6
site:EcoRI
end_sequence:BH305317"
244828..245823
/note="wgs end-extension
clone_end:"Sp6"
42200 c 42265 g 4973 t 33317 others
```

QY	1	AATGTCCTAAGCCTTTT	19	95.0%;	Score 19;	DB 2;	Length 248752;
				Best Local Similarity	100.0%;	Pred. No. 1.1e+02;	
				Matches 19;	Conservative 0;	Mismatches 0;	Indels 0;
							Gaps 0;
DB	203668	AATGTCCTAAGCCTTTT	203650				

RESULT 11	AC121457	LOCUS	DEFINITION	AC121457	264615 bp	DNA	linear	HTG 12-OCT-2002
AC121457			Rattus norvegicus clone CH230-146J5, *** SEQUENCING IN PROGRESS					
AC121457			***, 2 unordered pieces.					
AC121457			AC121457					
AC121457.3			GI:23907899					
HTG; HTGS_PHASE1			HTGS_DRAFT; HTGS_ENRICHED.					
Rattus norvegicus			Rattus norvegicus (Norway rat)					
Rattus norvegicus								
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.								
1 (baes 1 to 264615)								
REFERENCE								
1								
AUTHORS								
Muzny D, Marie M								

REFERENCE  
AUTHORS

1 (bases 1 to 264615)  
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguitano, D.,

TITLE	TITLE
JOURNAL	Direct Submission
REFERENCE	2 (bases 1 to 264615)
AUTHORS	Morley, R.C.
TITLE	Direct Submission
JOURNAL	Submitted (18-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 264615)
AUTHORS	Rat Genome Sequencing Consortium.
TITLE	Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
JOURNAL	On Oct 12, 2002 this
COMMENT	

Anyalabechi V., Aoyagi, A., Ayodeji, M., Bacs, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnesread, M., Benahmed, F., Biant, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Bunney, C., Burch, C., Cavazos, I., Cestari, K., Calderon, E., Chacko, J., Carter, K., Cavazos, I., Cestari, H., Center, A., Chavel, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Davila, M. L., Davis, C., Davy-Carroll, L., De Ande, A., D'Souza, L., Delgado, O., Benson, S., Detamco, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, C., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabris, A., Gante, R., Garcia, A., Garner, T., Gazda, M., Gebregziabir, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howell, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, J., Jackson, L., Jacob, L., Jiang, X., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Koyar, C., Kowals, C., Kraft, C. L., Ledow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, M., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenzen, K., L., Louised, H., Lozada, R. J., Lu, X., Ma, U., Maheshwari, L., Mahindaratne, M., Mahmud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mihaljevic, A., Miner, G., Minna, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Muntalasa, M., Murphy, M., Naik, L., Naeckelmeier, O., Okwom, G., Olarunpusgoun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, R., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., A., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Sanders, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J., Shetty, J., Shvatsky, E., A., Sisson, I., Sitter, C. D., Smales, D., Sneed, A., Soderstrom, A., Song, X., Z., Sorrell, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svalok, A., Taber, P., Taylor, J., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., Weller, P., Williams, G., Willson, R., Wleciyk, R., Wooden, H., Wotley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, Y., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Weidenhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Direct Submission

2. (bases 1 to 264615)

Unpublished

2. (bases 1 to 264615)

Worley, K.C.

Direct Submission

Submitted (18-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3. (bases 1 to 264615)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (12-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Oct 12, 2002 this sequence version replaced gi:21908060.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature



## table.

## ----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information

Center project name: GYKE

Center clone name: CH230-146J5

----- Summary Statistics

Assembly program: PHRAP, version 0.990329

Consensus quality: 207625 bases at least Q40

Consensus quality: 211034 bases at least Q30

Consensus quality: 213357 bases at least Q20

Estimated insert size: 215749; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length

\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 2 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 263260: contig of 263260 bp in length

\* 263261: 263360: gap of unknown length

\* 263361: 264615: contig of 1255 bp in length.

\* Location/Qualifiers

1..264615

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-146J5"

1..1514

/note="wgs contig"

misc\_feature

1565..3140

/note="wgs contig"

misc\_feature

4867..6350

/note="wgs contig"

misc\_feature

complement(262294..263119)

/note="clone boundary"

clone\_end:77

site:BCCK1

end\_sequence:RWBAG51TJB"

BASE COUNT 63778 a 42319 c 42274 g 67234 t 49010 others

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 19; DB 2; Length 264615;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AATGTCCTTAAGCTTTT 19

11111111111111111111

Db 97396 AATGTCCTTAAGCTTTT 97414

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

11111111111111111111

## TITLE

Miki, T. and Brandi, M.L.  
Calcium modulates the cyclin D1 expression in a rat parathyroid cell line

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

Brandi, M. L.  
Submitted (23-SEP-1993) M. Brandi, Endocrine Unit, University of Florence, Dept of Clinical Pathology, Viale Piseracini 6, 50139 Firenze, ITALY

Location/Qualifiers

1..1454

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/db\_xref="taxon:10116"

/cell\_line="patented rat parathyroid cell line"

1..1454

/gene="CCND1"

153..1040

/gene="CCND1"

/codon\_start=1

/product="CYCLIN D1"

/protein\_id="CA53020.1"

/protein\_id="CA53020.1"

/db\_xref="GI:473123"

/db\_xref="SWISS-PROT:P39948"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

/db\_xref="EMBL:U00001.1"

Schizosaccharomyces pombe (fission yeast)

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

Schizosaccharomyces pombe

BASE COUNT	611 a	587 c	617 g	543 t
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	1	1
12	1	1	1	1
13	1	1	1	1
14	1	1	1	1
15	1	1	1	1
16	1	1	1	1
17	1	1	1	1
18	1	1	1	1
19	1	1	1	1
20	1	1	1	1
21	1	1	1	1
22	1	1	1	1
23	1	1	1	1
24	1	1	1	1
25	1	1	1	1
26	1	1	1	1
27	1	1	1	1
28	1	1	1	1
29	1	1	1	1
30	1	1	1	1
31	1	1	1	1
32	1	1	1	1
33	1	1	1	1
34	1	1	1	1
35	1	1	1	1
36	1	1	1	1
37	1	1	1	1
38	1	1	1	1
39	1	1	1	1
40	1	1	1	1
41	1	1	1	1
42	1	1	1	1
43	1	1	1	1
44	1	1	1	1
45	1	1	1	1
46	1	1	1	1
47	1	1	1	1
48	1	1	1	1
49	1	1	1	1
50	1	1	1	1
51	1	1	1	1
52	1	1	1	1
53	1	1	1	1
54	1	1	1	1
55	1	1	1	1
56	1	1	1	1
57	1	1	1	1
58	1	1	1	1
59	1	1	1	1
60	1	1	1	1
61	1	1	1	1
62	1	1	1	1
63	1	1	1	1
64	1	1	1	1
65	1	1	1	1
66	1	1	1	1
67	1	1	1	1
68	1	1	1	1
69	1	1	1	1
70	1	1	1	1
71	1	1	1	1
72	1	1	1	1
73	1	1	1	1
74	1	1	1	1
75	1	1	1	1
76	1	1	1	1
77	1	1	1	1
78	1	1	1	1
79	1	1	1	1
80	1	1	1	1
81	1	1	1	1
82	1	1	1	1
83	1	1	1	1
84	1	1	1	1
85	1	1	1	1
86	1	1	1	1
87	1	1	1	1
88	1	1	1	1
89	1	1	1	1
90	1	1	1	1
91	1	1	1	1
92	1	1	1	1
93	1	1	1	1
94	1	1	1	1

Search completed: August 1, 2003, 17:32:55  
Job time : 84.1573 secs

1. .2358

Sun Aug 3 09:03:58 2003

us-09-834-291-14.rst

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:13:22 ; Search time 68.4754 Seconds  
(without alignments)  
7098.748 Million cell updates/sec

Title: US-09-834-291-14  
Perfect score: 20  
Sequence: 1 aatgtgcttaagctttttt 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: em\_esta:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_escov:\*
- 6: em\_escpl:\*
- 7: em\_escro:\*
- 8: em\_hic:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_esc5:\*
- 15: em\_escfun:\*
- 16: em\_escrom:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_pln:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_mam:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rtd:\*
- 26: em\_gss\_phg:\*
- 27: em\_gss\_vrl:\*
- 28: gb\_gss1:\*
- 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	95.0	468	28	BH867204 hg93f08.y
2	19	95.0	617	29	BZ674215 PUBIT774D
3	18.4	92.0	159	9	AV243724 AV243724
4	18.4	92.0	160	9	AV174680 AV174680

5	18.4	92.0	198	9	AW213875	AW213875 uc044d03.x
6	18.4	92.0	213	9	AV134425	AV134425 AV134425
7	18.4	92.0	226	10	BH834512	BH834512 BH834512
8	18.4	92.0	237	9	AV133072	AV133072 AV133072
9	18.4	92.0	240	10	BH048825	BH048825 BH048825
10	18.4	92.0	255	10	BH546588	BH546588 BH546588
11	18.4	92.0	270	9	AV135706	AV135706 AV135706
12	18.4	92.0	278	12	B1289904	B1289904 UI-R-DKO-
13	18.4	92.0	280	10	BH049315	BH049315 BH049315
14	18.4	92.0	283	10	BF323470	BF323470 uc282b07.x
15	18.4	92.0	283	9	AV219191	AV219191 AV219191
16	18.4	92.0	299	9	AV032131	AV032131 AV032131
17	18.4	92.0	328	28	AQ072724	AQ072724 HS_2255_A
18	18.4	92.0	332	9	AW229904	AW229904 uc044d03.y
19	18.4	92.0	339	10	BH546854	BH546854 BH546854
20	18.4	92.0	379	13	W78658	W78658 me88b07.r1
21	18.4	92.0	379	13	BY229322	BY229322 BY229322
22	18.4	92.0	380	13	BY174614	BY174614 BY174614
23	18.4	92.0	388	10	BH836659	BH836659 BH836659
24	18.4	92.0	411	10	BH781889	BH781889 BH781889
25	18.4	92.0	444	12	B130131	B130131 UI-R-DQO-
26	18.4	92.0	446	13	BY472560	BY472560 BY472560
27	18.4	92.0	457	10	BH831513	BH831513 BH831513
28	18.4	92.0	469	12	BH383441	BH383441 UI-R-DML-
29	18.4	92.0	470	12	BH898746	BH898746 UI-M-DJ0-
30	18.4	92.0	491	14	CA877820	CA877820 K0960E07-
31	18.4	92.0	515	12	B1452813	B1452813 603170340
32	18.4	92.0	515	14	CA879995	CA879995 K0979F03-
33	18.4	92.0	516	14	CA872437	CA872437 K0917H06-
34	18.4	92.0	517	14	CA870300	CA870300 K0910A03-
35	18.4	92.0	518	14	CA870688	CA870688 K0904C02-
36	18.4	92.0	518	14	CA871824	CA871824 K0913E03-
37	18.4	92.0	518	14	CA873313	CA873313 K0924C08-
38	18.4	92.0	527	12	CA879757	CA879757 K0976E11-
39	18.4	92.0	529	9	BM932381	BM932381 UI-M-BH2-
40	18.4	92.0	538	14	BM925960	BM925960 um79h05.y
41	18.4	92.0	538	10	BE979816	BE979816 UI-M-BG2-
42	18.4	92.0	570	14	CA873049	CA873049 K0922E01-
43	18.4	92.0	574	12	BM386791	BM386791 UI-R-CN1-
44	18.4	92.0	577	28	AZ409434	AZ409434 1M0181X04
45	18.4	92.0	578	13	BQ510695	BQ510695 EST618110

ALIGNMENTS

RESULT 1  
LOCUS BH867204 468 bp DNA linear GSS 05-AUG-2002  
DEFINITION hg93f08.y9 WGS-Zmap9f (UM107 adapted methyl filtered) Zea mays  
genomic clone hg93f08 5', genomic survey sequence.

ACCESSION BH867204  
VERSION BH867204.1 GI:22103101  
KEYWORDS GSS  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 468)  
Rabinowicz, P.D., O'Shaughnessy, A.L., Ballja, V., Dedhia, N.,  
Katzemburger, F., King, L., Miller, B., Muller, S.,  
Zutavern, T., McCombie, W.R., and Martienssen, R.A.  
Genomic shotgun sequences from Zea mays (methyl-filtered)  
Unpublished  
Contact: W. Richard McCombie  
Lila Annenberg Hazen Genome Sequencing Center  
Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel.: 516 367 8884  
Fax: 516 367 8874  
Email: mccombie@cshl.org  
Plate: hg93 row: F column: 08

TITLE  
JOURNAL  
COMMENT

Seq primer: -21M13UnivRev  
 Class: shotgun  
 High quality sequence stop: 468.  
 Location/Qualifiers

## FEATURES

1. 468

/organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /cultivar="B73"  
 /db\_xref="taxon:4577"  
 /clone="hg93f08"

/lab\_host="JM107 or DH5a"  
 /note="Organ: Immature ears; site 1: Xba I; site 2: Xba I; added by fill in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were ligated into M13mp19, b/g reads in pUC19. The same ligation was transformed in either JM107 or DH5a."

BASE COUNT  
 119 a 79 c 95 g 175 t

Query Match  
 Best Local Similarity 95.0%; Score 19; DB 28; Length 468;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ATGTTGCTTAAGCTTTT 20  
 300 ATGTTGCTTAAGCTTTT 318

RESULT 2  
 B2674215 617 bp DNA linear GSS 05-FEB-2003  
 LOCUS PUBIT74TD\_ZM\_0.6-1.0\_KB Zea mays genomic clone ZMMB7A064N04,  
 DEFINITION genomic survey sequence.  
 ACCESSION B2674215  
 VERSION B2674215.1 GI:28224015  
 KEYWORDS GSS.

SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 1 (bases 1 to 617)  
 White, C.A., Quackenbush, J., Van Aken, S., Utechtack, T., Resnick,  
 A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennetzen, J.  
 Maize Genomics Consortium  
 Unpublished  
 Contact: Cathy Whitelaw

REFERENCE  
 TITLE 9712 Medical Center Drive, Rockville, MD 20850, USA  
 JOURNAL Tel: 301-838-5843  
 COMMENT Fax: 301-838-0208  
 Email: whitelaw@igr.org  
 Seq primer: TP  
 Class: sheared ends.

FEATURES  
 source Location/Qualifiers

1. 617  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMB7A064N04"  
 /clone\_1lb="ZM\_0.6-1.0\_KB"  
 /note="Vector: PCR4-TOP0; Site 1: EcoRI; 0.6-1.0 kb high  
 Cor selected genomic DNA library"  
 BASE COUNT 166 a 103 c 235 t

Query Match  
 Best Local Similarity 95.0%; Score 19; DB 29; Length 617;  
 100.0%; Pred. No. 1.1e+03;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 ATGTTGCTTAAGCTTTT 20  
 519 ATGTTGCTTAAGCTTTT 537

RESULT 3  
 AV243724 159 bp mRNA linear EST 04-NOV-1999  
 LOCUS AV243724 RIKEN full-length enriched, 0 day neonate head Mus  
 DEFINITION musculus cDNA clone 4831423J17 3' similar to S78355 Cyl-1-cyclin D1  
 mRNA sequence.  
 ACCESSION AV243724  
 VERSION AV243724.1 GI:6231183  
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 AUTHORS Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T.,  
 Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F.,  
 Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai,  
 C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M.,  
 Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,  
 Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata,  
 Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H.,  
 Suzuki, H., Takahashi, F., Tateo, M., Tomihata, N., Tsunoda, Y.,  
 Watanabe, A., Watanabe, S., Yamamura, T., Yasuniishi, A., Yokota, T.,  
 Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Kono, H., et al. 1999)  
 Unpublished

COMMENT  
 TITLE Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 JOURNAL Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gs.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Sasaki, N., Izawa, M., Watanabe, M., Okazaki, Y. and Hayashizaki,  
 Y., Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki,  
 Y.  
 'Transcriptional sequencing: A method for DNA sequencing using RNA  
 polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)  
 Itoh, M., Katsunari, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, D.,  
 Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,  
 Y. and Hayashizaki, Y.  
 Automated filtration-based high-throughput plasmid preparation  
 system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
 19-44 (1999)  
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for  
 further details.

FEATURES  
 source Location/Qualifiers

1. 159  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="4831423J17"  
 /sex="mixed"  
 /tissue\_type="head"  
 /dev\_stage="0 day neonate"  
 /lab\_host="DHI0B"  
 /clone\_1lb="RIKEN full-length enriched, 0 day neonate  
 head"  
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken



DEFINITION AV134425 Mus musculus C57BL/6J 10-11 day embryo Mus musculus cDNA  
 AV134425  
 AV134425 GI:5320155  
 EST.  
 Mus musculus (house mouse)  
 SOURCE  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,  
 Akahira, S., Akiyama, Y., Fukuda, S., Fukunishi, Y., Funayama, T., Hara  
 Kikuchi, N., Kojima, Y., Matsuyama, T., Nishimura, H., Oda, H., Owa, C.,  
 Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara  
 Y., Suzuki, H., Suzuki, H., Tachino, M., Tomaru, Y., Tomihaga, N.,  
 Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
 RIKEN Mouse ESTs  
 Unpublished  
 JOURNAL  
 COMMENT  
 Contact: Chie Owa  
 Genome Science Laboratory  
 RIKEN  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-9145  
 Fax: 81-298-36-9098  
 Email: genome-res@rtc.riken.go.jp  
 Thermolabile and thermostable activation of thermolabile enzymes by  
 trehalose and its application for the synthesis of full length cDNAs by  
 Transcriptional sequencing: A method for DNA sequencing using RNA  
 polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3435-3460 (1998))  
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for  
 further details.

FEATURES  
 source  
 1. 213  
 /location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="2810006L11"  
 /sex="mixed"  
 /dev\_stage="10-11 day embryo"  
 /clone\_id="Mus musculus C57BL/6J 10-11 day embryo"  
 /clone\_id="Mus musculus C57BL/6J 10-11 day embryo"  
 BASE COUNT 61 a 22 c 54 g 76 t  
 ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 213;  
 Best Local Similarity 95.0%; Pred. No. 1.5e+03;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AATGTCCTTAAGCTTTT 20  
 Db 208 AATGTCCTTAAGCTTTT 189

RESULT 7  
 BB834512/c  
 LOCUS BB834512 226 bp mRNA linear EST 19-NOV-2001  
 DEFINITION BB834512 RIKEN full-length enriched, mammary gland RCB-0527  
 JYG-MC(B) cDNA Mus musculus cDNA clone G930031G12 3', mRNA  
 accession BB834512  
 version BB834512  
 keywords BB834512.1 GI:17012755  
 EST.  
 Mus musculus (house mouse)  
 SOURCE  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Carninci, P., Aizawa, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,  
 Akahira, S., Akiyama, Y., Fukuda, S., Fukunishi, Y., Funayama, T., Hara  
 Kikuchi, N., Kojima, Y., Matsuyama, T., Nishimura, H., Oda, H., Owa, C.,  
 Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara  
 Y., Suzuki, H., Suzuki, H., Tachino, M., Tomaru, Y., Tomihaga, N.,  
 Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
 RIKEN Mouse ESTs  
 Unpublished  
 JOURNAL  
 COMMENT  
 Contact: Chie Owa  
 Genome Science Laboratory  
 RIKEN  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-9145  
 Fax: 81-298-36-9098  
 Email: genome-res@rtc.riken.go.jp  
 Thermolabile and thermostable activation of thermolabile enzymes by  
 trehalose and its application for the synthesis of full length cDNAs by  
 Transcriptional sequencing: A method for DNA sequencing using RNA  
 polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3435-3460 (1998))  
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for  
 further details.

TITLE  
 JOURNAL  
 COMMENT  
 Contact: Yoshinobu Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9226  
 Email: genome-res@rtc.riken.go.jp  
 URL: <http://genome.gsc.riken.go.jp/>  
 Carninci, P., Shibata, Y., Hayashizaki, Y., Sugahara, Y., Shibata, K., Itoh  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 Wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iizawa, M., Ohara, E.,  
 Sato, K., Kawai, J., Okazaki, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama  
 Hayashizaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
 Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.  
 e mouse tissues.

FEATURES  
 source  
 1. 226  
 /location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="G930031G12"  
 /issue\_type="mammary gland"  
 /cell\_line="RCB-0527 JYG-MC(B)"  
 /clone\_id="RCB-0527 JYG-MC(B) cDNA"  
 /clone\_id="RCB-0527 JYG-MC(B) cDNA"  
 BASE COUNT 81 a 31 c 48 g 66 t  
 ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 226;  
 Best Local Similarity 95.0%; Pred. No. 1.5e+03;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AATGTCCTTAAGCTTTT 20  
 Db 210 AATGTCCTTAAGCTTTT 191

RESULT 8  
 AV133072/c  
 LOCUS AV133072 237 bp mRNA linear EST 01-JUL-1999  
 DEFINITION AV133072 Mus musculus C57BL/6J 11-day embryo Mus musculus cDNA  
 accession AV133072  
 version AV133072  
 keywords AV133072.1 GI:5308802  
 EST.  
 Mus musculus (house mouse)  
 SOURCE  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,  
 Akahira, S., Akiyama, Y., Fukuda, S., Fukunishi, Y., Funayama, T., Hara  
 Kikuchi, N., Kojima, Y., Matsuyama, T., Nishimura, H., Oda, H., Owa, C.,  
 Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara  
 Y., Suzuki, H., Suzuki, H., Tachino, M., Tomaru, Y., Tomihaga, N.,  
 Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
 RIKEN Mouse ESTs  
 Unpublished  
 JOURNAL  
 COMMENT  
 Contact: Yoshinobu Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9226  
 Email: genome-res@rtc.riken.go.jp  
 URL: <http://genome.gsc.riken.go.jp/>  
 Carninci, P., Shibata, Y., Hayashizaki, Y., Sugahara, Y., Shibata, K., Itoh  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 Wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iizawa, M., Ohara, E.,  
 Sato, K., Kawai, J., Okazaki, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama  
 Hayashizaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
 Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.  
 e mouse tissues.

TITLE  
JOURNAL  
COMMENT

Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara, A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Naitsumu, H., Oda, H., Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, T., Sugahara, Y., Suzuki, H., Suzuki, H., Tateo, M., Tomaru, Y., Tomioka, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.

RIKEN Mouse ESTs  
Unpublished  
Contact: Chile Owa  
Genome Science Laboratory  
RIKEN  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-9145  
Fax: 81-298-36-9098

Email: genome-res@rtc.riken.go.jp  
Thermostabilization and thermoinactivation of thermostable enzymes by trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))  
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES  
source

Location/Qualifiers  
1..237  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="270095E14"  
/sex="mixed"  
/dev\_stage="11-day embryo"  
/clone\_11b="Mus musculus C57BL/6J 11-day embryo"

BASE COUNT  
ORIGIN

83 a 27 c 49 g 78 t  
Query Match 92.0%; Score 18.4; DB 9; Length 237;  
Best Local Similarity 95.0%; Pred. No. 1.6e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 AATGTTGCTTAAGCTTTT 20  
Db 232 AATGTTGCTTAAGCTTTT 213

RESULT 9  
BB048825/c  
LOCUS  
DEFINITION  
BB048825 RIKEN full-length enriched, adult male olfactory bulb Mus musculus cDNA clone 6430596C05 3' similar to S78355 Cyl-1-cyclin D1 mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BB048825  
BB048825.1 GI:8455973  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 240)  
Komno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hitozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koyama, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, Y., Sugahara, Y., Y. Shigemoto, Y., Shiraki, T., Sogabe, T., Tomioka, N., Toyama, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., T. Tanoda, Y., Watanabe, S., Watanabe, S., Yamamura, T., Yamazaki, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

TITLE  
JOURNAL  
Unpublished

## COMMENT

Contact: Yoshinori Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-72 Saito-cho, Tsukuba, Ibaraki, 305-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@rtc.riken.go.jp  
URL: <http://genome.gsc.riken.go.jp/>  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
Thermostabilization and thermoinactivation of thermostable enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998)  
CDNA. Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998)  
Itoh, M., Kikuchi, N., Akiyama, J., Shibata, Y., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9(5):463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES  
source

Location/Qualifiers  
1..240  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="6430596C05"  
/sex="male"  
/tissue\_type="olfactory brain"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_11b="RIKEN full-length enriched, adult male olfactory bulb"

BASE COUNT  
ORIGIN

88 a 24 c 48 g 80 t  
Query Match 92.0%; Score 18.4; DB 10; Length 240;  
Best Local Similarity 95.0%; Pred. No. 1.6e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 AATGTTGCTTAAGCTTTT 20  
Db 233 AATGTTGCTTAAGCTTTT 214

RESULT 10  
BB546588/c  
LOCUS  
DEFINITION  
BB546588 RIKEN full-length enriched, 0 day neonate eyeall Mus musculus cDNA clone E130310C20 3' similar to S78355 Cyl-1-cyclin D1 mRNA sequence.

ACCESSION  
BBS46588  
VERSION  
BBS46588.1 GI:9618016  
KEYWORDS  
EST  
SOURCE  
MUS MUSCULUS (house mouse)

REFERENCE  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 255)

AUTHORS  
Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci  
Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,  
Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,  
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,  
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,  
Oto, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y.,  
Shigemoto, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y.,  
Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya,  
T., Tanoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yamataka, I.,  
Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino,  
M., Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Kono, H., et al.)  
Unpublished

TITLE  
JOURNAL  
COMMENT  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.go.jp  
URL: http://genome.gsc.riken.go.jp/  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki,  
N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Thermofabilitation and thermocactivation of thermolabile enzymes by  
trehalose and its application for the synthesis of full length  
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh, M., Kikuchi, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,  
Y. and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation  
System. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High efficiency full-length cDNA cloning. Methods Enzymol. 303,  
19-44 (1999)  
Please visit our web site (http://genome.riken.go.jp) for  
further details.

FEATURES  
source  
location/Qualifiers

1..255  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="E130310020"  
/issue\_type="eyeball"  
/dev\_stage="0 day neonate"  
/lab\_host="DH10B"  
/clone\_1lb="RIKEN full-length enriched, 0 day neonate  
eyeball"  
/note="Site 1: Sali; Site 2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer 15'  
GAGAGAGAGCGCCGCACTGAGTCTTTTCTTTTCTTTT 3'. cDNA was  
transcribed using trehalose thermo-activated reverse  
cap-trapper. Second strand cDNA was prepared with the  
primer adapter of sequence 15'  
GAGAGAGAGTCTCGAGTTATTAATTAATCCCTCCCTCCCT 3'. cDNA  
was cleaved with BamHI and XhoI. Vector: a modified  
pBluescript KS(+) after bulk excision from lambda FLIC I."

BASE COUNT  
89 a 31 c 57 g 78 t

## ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 255;  
Best Local Similarity 95.0%; Pred. No. 1.6e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Db 1 AATGTTGCTTAAGCTTTT 20  
248 AATGTTGCTTAAGCTTTT 229

## RESULT 11

AV135706/c  
LOCUS  
DEFINITION  
AV135706 Mus musculus C57BL/6J 10-11 day embryo Mus musculus cDNA  
clone 2810020K08, mRNA sequence.  
AV135706  
ACCESSION  
AV135706.1 GI:5321436  
VERSION  
KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)  
Mus musculus

ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 270)  
REFERENCE  
Carninci, P., Shibata, K., Ozawa, Y., Kono, H., Itoh, M., Aizawa, K.,  
A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J.,  
Kikuchi, N., Kojima, Y., Matsuyama, T., Nishiyama, H., Oda, H., Owa, C.,  
Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara,  
Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya,  
T., Tanoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yokota, T., Yoshida, K.,  
Yoshiki, A., Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
RIKEN Mouse ESTs  
Unpublished

TITLE  
JOURNAL  
COMMENT  
Contact: Chie Owa  
Genome Science Laboratory  
RIKEN  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-9145  
Fax: 81-298-36-9098  
Email: genome-res@sc.riken.go.jp  
Thermofabilitation and thermocactivation of thermolabile enzymes by  
trehalose and its application for the synthesis of full length cDNA  
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))  
Transcriptional sequencing: A method for DNA sequencing using RNA  
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))  
Please visit our web site (http://genome.riken.go.jp) for  
further details.

## FEATURES

source

location/Qualifiers

1..270  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="2810020K08"  
/sex="mixed"  
/dev\_stage="10-11 day embryo"  
/clone\_1lb="Mus musculus C57BL/6J 10-11 day embryo"  
BASE COUNT  
88 a 34 c 61 g 87 t

## BASE COUNT

ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 270;  
Best Local Similarity 95.0%; Pred. No. 1.6e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

## Db

1 AATGTTGCTTAAGCTTTT 20  
265 AATGTTGCTTAAGCTTTT 246

RESULT 12  
BI289904/c  
LOCUS  
BI289904 278 bp mRNA linear EST 19-JUL-2001



DEFINITION UI-R-DKO-cfp-f-04-0-UI.s1 UI-R-DKO Rattus norvegicus cDNA clone  
 ACCESSION UI-R-DKO-cfp-f-04-0-UI 3', mRNA sequence.  
 B1289904  
 VERSION B1289904.1 GI:14947954  
 KEYWORDS EST.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 278)  
 Bonaldo, M.F., Lennon, G. and Soares, M.B.  
 Normalization and subtraction: two approaches to facilitate gene discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8889548  
 COMMENT Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Oligo-dt track not found, Not I site shown in beginning of sequence is likely internal to the message. cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)  
 Seq primer: M13 Forward  
 POLYN=No.

# FEATURES

Source

1. 278  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UI-R-DKO-cfp-f-04-0-UI"  
 /dev\_stage="ADULT"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_libs="UI-R-DKO"  
 /note="Vector: pRT73D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; The UI-R-DKO library is a subtracted library derived from a mixture of five individually tagged normalized rat libraries: brain-nRBP (20%), heart-nRBP (20%), kidney-nRBP (20%), aorta-nRBP (20%), and placenta-nRBP (20%). Each original library was constructed from a mixture of equal amounts of RNA from seven different developmental time-points: embryonic day 17, embryonic day 19, embryonic day 21, adult day 1, adult day 12, adult day 75, and adult day 200. (Exception: the aorta pool does not contain embryonic day 17 RNA and the placenta pool contains only the three embryonic stages). Each library was normalized individually according to the procedure described by Bonaldo, Lennon & Soares (Genome Research 6: 791-806, 1996). For construction of the DKO subtracted library, plasmid DNA from each of the five individually tagged normalized libraries was mixed in the proportions specified above and electroporated into competent bacteria for production of single-stranded circular DNA representing the pool of libraries. Single-stranded circular DNA representing these five normalized libraries was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plasmid DNA template preparation) comprising: a) a set of about 1,000 arrayed clones from each of the five non-normalized libraries of brain (CT08), heart (CS08), kidney (CU08), aorta (CW08), and placenta (CX08). The resulting pool of approximately 5,000 clones represented about 33.3% of the final driver population. A set of about 2,000 arrayed clones from each of the five normalized libraries of brain (CT0), heart (CS0), kidney (CU0), aorta (CW0), and placenta (CX0). The resulting pool of about 10,000 clones

BASE COUNT 107 a 41 c 58 g 72 t  
 ORIGIN  
 Query Match 92.0%; Score 18.4; DB 12; Length 278;  
 Best Local Similarity 95.0%; Pred. No. 1.6e+03;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 AATGTTGCTTAAGCTTTT 20  
 |||||  
 Db 247 AATGTTTCTTAAGCTTTT 228

RESULT 13  
 BB049315/c  
 LOCUS  
 DEFINITION BB049315 280 bp mRNA linear EST 25-JUN-2000  
 BB049315 RIKEN full-length enriched, adult male cerebellum Mus musculus cDNA clone 6530404H10 3' similar to S78355 Cyl-1-cyclin D1 mRNA sequence.  
 ACCESSION BB049315 GI:8456463  
 VERSION BB049315  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 280)  
 Konno, H., Aizawa, K., Akahira, S., Akiyama, D., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozawa, T., Horii, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya, T., Tsunoda, Y., Watanabe, A., Watanabe, S., Yamamura, T., Yamataka, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Konno, H., et al.)  
 Unpublished

TITLE Unpublished  
 JOURNAL  
 COMMENT Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Saito-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@sc.riken.go.jp/  
 URL: http://genome.res.riken.go.jp/  
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 The establishment and thermoactivation of the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh, M., Katsunari, T., Akiyama, U., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
 Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

# FEATURES

Source

1. 280  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="6530404H10"

```

/sex="male"
/tissue_type="cerebellum"
/dev stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male
cerebellum"
/notes="Site 1: XhoI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'].
GAGAGAGAGAGATCCAAAGAGCTCTTTTTTTTTTTTTTTT 3', cDNA was
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'].
GAGAGAGAGAGCGCGCCGCAATTAATTCGATTAATTAATTAATCCCCCCC
3']. cDNA was cloned into the XhoI and BamHI sites.
Vector: a modified phluescript KS(+) after bulk excision
from lambda FICIT"
BASE COUNT
90 a 40 c 59 g 91 t
ORIGIN
Query Match 92.0%; Score 18.4; DB 10; Length 280;
Best Local Similarity 95.0%; Pred. No. 1.6e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY
1 AATGTTGCTTAAGCTTTT 20
|||||
266 AATGTTCTTAAAGCTTTT 247
RESULT 14
LOCUS
BF323470 283 bp mRNA linear EST 29-DEC-2000
DEFINITION
u82b07.x1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3675541 3',
mRNA sequence.
ACCESSION
BF323470
VERSION
BF323470.1 GI:11273070
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 283)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished
COMMENT
Other ESTs: u82b07.y1
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
image.lnl.gov/image/html/lresources.shtml
MGI:1436309
Seg primer: -40UP from Gibco.
FEATURES
Location/Qualifiers
1..283
/mol_type="mRNA"
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:3675541"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."

```

```

/lab_host="DH10B"
/clone_lib="NCI_CGAP Lu29"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT
87 a 57 c 40 g 99 t
ORIGIN
Query Match 92.0%; Score 18.4; DB 10; Length 283;
Best Local Similarity 95.0%; Pred. No. 1.6e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY
1 AATGTTGCTTAAGCTTTT 20
|||||
10 AATGTTCTTAAAGCTTTT 29
RESULT 15
LOCUS
AV219191/c 298 bp mRNA linear EST 01-NOV-1999
DEFINITION
AV219191 RIKEN full-length enriched, 12 days embryo head Mus
musculus cDNA clone 3010086E05.3, similar to S78355 Cyl-1-cyclin D1
[ mice, BALB/c, brain, mRNA, 3737 nt], mRNA sequence.
ACCESSION
AV219191
VERSION
AV219191.1 GI:6168368
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 298)
Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T.,
Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F.,
Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai
C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M.,
Matsumura, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata
Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H.,
Suzuki, H., Takehashi, F., Tateo, M., Tomimaga, N., Tsunoda, Y.,
Watahiki, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T.,
Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Kono, H., et al. 1999)
JOURNAL
Unpublished
TITLE
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Sasaki, N., Izawa, M., Matshiki, M., Ozawa, K., Tanaka, T., Yoneda, Y.,
Matsumura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki
Y.
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
Itoh, M., Kikunishi, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki
Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.
FEATURES
Location/Qualifiers
1..298
/mol_type="mRNA"
/organism="Mus musculus"
/strain="C57BL/6J"

```

```

/db_xref="taxon:10090"
/clone="3010086E05"
/sex="mixed"
/tissue_type="head"
/dev_stage="12 days embryo"
/lab_host="SOLR"
/clone_1ib="RIKEN full-length enriched, 12 days embryo
head"
/note="Site 1: XhoI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTVA 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 50.0. Second strand
cDNA was prepared with the primer/adaptor of sequence [5'
GAGAGAGAGATTCGAGTTAATTAATATCCCCCCCCC 3']"
BASE COUNT      93 a      48 c      51 g      106 t
ORIGIN

```

```

Query Match      92.0%; Score 18.4; DB 9; Length 298;
Best Local Similarity 95.0%; Pred. No. 1.6e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 AATGTTGCTTACGCTTTT 20
      |||||  |||||  |||||  |||||
DB      293 AATGTTCTTACGCTTTT 274

```

Search completed: August 1, 2003, 20:43:06  
 Job time : 71.4754 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:32:17 ; Search time 6.69868 Seconds  
(without alignments)  
8059.612 Million cell updates/sec

Title: US-09-834-291-14  
Perfect score: 20  
Sequence: 1 aatgtgcttaagcttttt 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

N\_Geneseq\_19Jun03:\*

1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*
25:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	20	100.0	21	AAZ88724
2	18.4	92.0	22	ABN79875
3	16.8	84.0	24	AAFP3265
4	16.8	84.0	21	AAAP7038
5	16.8	84.0	21	AAAS2136
6	16.8	84.0	23	AAAP7060
7	16.8	84.0	23	ABLI4642
8	16.8	84.0	24	ABT10146

9	16.8	84.0	154902	24	ABQ88198
10	16.8	84.0	165199	22	ABX83460
11	16.4	82.0	333	22	AA860064
12	16.4	82.0	564	22	AAFP3687
13	16.4	82.0	1404	21	AAAC5476
14	16.4	82.0	2613	23	ABLI7720
15	16.4	82.0	3945	23	ABLI22326
16	16.4	82.0	30610	22	ABLI5643
17	16.4	82.0	1503841	24	ABT00010
18	16.4	82.0	1503841	24	ABT01503
19	16.4	82.0	1503900	22	AAK95240
20	16.4	82.0	1503900	22	AAK96733
21	16.4	80.0	1537	21	AA45667
22	16.4	80.0	1544	21	AA436610
23	16.4	80.0	6165	21	ABN97371
24	15.8	79.0	289	22	AAK76919
25	15.8	79.0	289	22	AAK76920
26	15.8	79.0	289	22	AAK76922
27	15.8	79.0	289	22	AAK76922
28	15.8	79.0	289	22	AAK81702
29	15.8	79.0	289	22	AAK81704
30	15.8	79.0	300	21	AAAO0364
31	15.8	79.0	385	22	AA455104
32	15.8	79.0	413	22	AA185195
33	15.8	79.0	427	22	AA187437
34	15.8	79.0	461	22	AAFP3729
35	15.8	79.0	468	24	ABN96314
36	15.8	79.0	468	24	ABL67114
37	15.8	79.0	481	21	AA454504
38	15.8	79.0	513	22	AA452226
39	15.8	79.0	513	22	AAFP3441
40	15.8	79.0	533	24	ABK71753
41	15.8	79.0	628	22	AAFP3449
42	15.8	79.0	777	23	AA452154
43	15.8	79.0	792	23	AA452154
44	15.8	79.0	828	23	AA452154
45	15.8	79.0	885	24	ABK73291

#### ALIGNMENTS

RESULT 1  
AAZ88724  
AAZ88724 standard; DNA; 20 BP.

AAZ88724;  
11-MAY-2000 (first entry)

Human CD95 receptor promoter DNA p53 binding fragment #2.  
p53; CD95 receptor; human; screening; apoptosis-modulation;  
cancer chemotherapy; ss.

Homo sapiens.

DE19847779-Cl.  
03-FEB-2000.  
16-OCT-1998; 98DE-1047779.  
16-OCT-1998; 98DE-1047779.  
16-OCT-1998; 98DE-1047779.  
(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
Krammer P, Mueller-Schilling M, Oren M;  
WPI; 2000-162245/15.  
Novel receptor DNA useful for identifying apoptosis-modulating  
substances potentially useful for cancer chemotherapy

Human osteoblast d  
Human CDNA differe  
Human cancer agent  
CDNA encoding SRT  
Arabidopsis thalia  
Drosophila melanog  
Drosophila melanog  
Human nervous syst  
Human neuregulin 1  
Human neuregulin 1  
Human neuregulin-1  
Human neuregulin-1  
Arabidopsis thalia  
Arabidopsis thalia  
Gene #3869 used to  
Human immune/haema  
Human immune/haema  
Human immune/haema  
Human immune/haema  
Human immune/haema  
Human immune/haema  
Human colon cancer  
Novel human polynu  
Human polynucleoti  
CDNA encoding SRT  
Gene #2812 used to  
Thyroid cancer rel  
Arabidopsis thalia  
Arabidopsis thalia  
Aortic endothelial  
Human dthp polynu  
CDNA encoding SRT  
Staphylococcus aur  
Staphylococcus aur  
Bacillus lichenifo

XX Claim 2; Fig 5; 12pp; German.  
 XX  
 CC This invention describes a novel p53-binding region of a human CD95  
 CC receptor DNA molecule. The p53-binding region, or a vector containing  
 CC it, can be used to screen for apoptosis-modulating substances  
 CC potentially useful for cancer chemotherapy. This sequence represents a  
 CC fragment of the human CD95 receptor promoter which is capable of  
 CC binding p53.  
 CC  
 SQ Sequence 20 BP; 4 A; 2 C; 3 G; 11 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AATGTTGCTTAAGCTTTT 20  
 Db 1 AATGTTGCTTAAGCTTTT 20

RESULT 2  
 AAB79875/c  
 ID AAB79875 standard; DNA; 1689 BP.  
 XX  
 AC AAB79875;  
 XX  
 DT 24-JUL-2002 (first entry)  
 XX  
 DE Fungal ZBC gene sequence #112.  
 XX

Secondary metabolite; fungus; ZBC gene; zinc binuclear cluster protein;  
 KW antibacterial; beta-lactam; anti-hypercholesterolemic; lovastatin;  
 KW mevastatin; immunosuppressant; cyclosporin A; ergot alkaloid; ergocamine;  
 KW angiotensin inhibitor; ovalicin; glucan synthase inhibitor; gliotoxin;  
 KW fungal toxin; cell surface receptor; plant growth regulator; pigment;  
 KW insecticide; antineoplastic; gene; ds.  
 XX  
 OS Unidentified.  
 OS  
 PN WO200224865-A2.  
 XX  
 PD 28-MAR-2002.  
 XX  
 PF 19-SEP-2001; 2001WO-US29288.  
 XX  
 PR 19-SEP-2000; 2000US-233564P.  
 XX  
 PA (MICR-) MICROBIA INC.  
 XX  
 PI Holtzman D, Madden K, Maxon M, Sherman A;  
 XX  
 DR WPI; 2002-352005/38.  
 XX  
 DR P-PSDB; ABP35686.  
 XX  
 PT  
 PT  
 PT

New method for improving the production of a secondary metabolite e.g.  
 PT antineoplastic agent, ergot alkaloid from a fungus involves modulation  
 PT of the expression of at least one zinc binuclear cluster protein gene  
 PT  
 Claim 2; SEQ ID 235; 49pp + sequence listing; English.

The invention relates to improving the production of a secondary  
 CC metabolite by a fungus. This involves modulating the expression of at  
 CC least one ZBC (zinc binuclear cluster protein) gene in a manner to  
 CC improve the yield of the secondary metabolite. Methods of the invention  
 CC may be used for improving the production of the secondary metabolite e.g.  
 CC antibacterial (such as beta-lactam), an anti-hypercholesterolemic (such  
 CC as lovastatin or mevastatin), an immunosuppressant (such as cyclosporin A),  
 CC an ergot alkaloid (such as ergotamine), an angiotensin inhibitor (such  
 CC as ovalicin), a glucan synthase inhibitor, gliotoxin family of compounds,  
 CC a fungal toxin, a modulator of cell surface receptor signalling, a plant  
 CC growth regulator, a pigment, an insecticide, or an antineoplastic

CC compound. The method results in a decrease in fermentor run-time, a  
 CC decrease in the size of the fermentor required for the production of  
 CC equivalent amounts of the secondary metabolite, or a decrease in the  
 CC biomass required for the production, which translates into decreased  
 CC waste that must be handled in downstream processing. The sequences given  
 CC in records AAB79764-AAB79911 represent ZBC genes of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SQ Sequence 1689 BP; 508 A; 307 C; 323 G; 551 T; 0 other;

Query Match 92.0%; Score 18.4; DB 24; Length 1689;  
 Best Local Similarity 95.0%; Pred. No. 1.2e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AATGTTGCTTAAGCTTTT 20  
 Db 516 AATGTTGCTTAAGCTTTT 497

RESULT 3  
 AAF93265  
 ID AAF93265 standard; cDNA; 514 BP.  
 XX  
 AC AAF93265;  
 XX  
 DT 21-MAY-2001 (first entry)  
 XX  
 DE Small intestine tissue cDNA encoding SRT protein SEQ ID 86.  
 XX  
 DE Human; SRT; gene therapy; gene mapping; tissue typing; ss.  
 XX  
 KW Homo sapiens.  
 XX  
 OS Homo sapiens.  
 OS  
 PN WO200107611-A2.  
 XX  
 PD 01-FEB-2001.  
 XX  
 PF 21-JUL-2000; 2000WO-US20006.  
 XX  
 PR 26-JUL-1999; 99US-0145701.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Baker KP, Goddard A, Wood WI;  
 XX  
 DR WPI; 2001-112729/12.  
 XX  
 PT  
 PT  
 PT

New isolated nucleic acid molecule encoding a SRT polypeptide is useful  
 PT for production of recombinant SRT polypeptides, gene mapping,  
 PT diagnosing genetic disorders and for gene therapy -  
 PT  
 Claim 2; Fig 86; 663pp; English.

Sequences AAF93180 - AAF93743 represent polynucleotide sequences encoding  
 CC human SRT proteins. The cDNA sequences are isolated from various  
 CC different human tissue cDNA libraries. The invention relates to a method  
 CC for detecting cDNA encoding an SRT protein, a vector containing cDNA  
 CC encoding SRT, a host cell transformed with the vector, an isolated SRT  
 CC polypeptide, and an antibody which binds to SRT. The polynucleotide  
 CC sequence can be used in gene therapy and is useful in the recombinant  
 CC production of SRT polypeptides, as a hybridisation probe to screen  
 CC libraries to isolate cDNAs with sequence identity to SRT polypeptides, to  
 CC map the gene encoding the SRT polypeptides and analysing genetic  
 CC disorders, tissue typing and disease tissue detection. The SRT  
 CC polynucleotide sequences can be used in polymerase chain reaction,  
 CC screening for new therapeutic molecules and generation of antisense RNA  
 CC and DNA.

Sequence 514 BP; 129 A; 111 C; 118 G; 152 T; 4 other;  
 Query Match 84.0%; Score 16.8; DB 22; Length 514;

Best Local Similarity 90.0%; Pred. No. 5.6e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AATGTTGCTTAAGCTTTT 20  
| | | | | | | | | |  
Db 78 AATGTTGCTGAAGCTTTCT 97

RESULT 4  
AAA97038  
ID AAA97038 standard; DNA; 1404 BP.

AC AAA97038;

DT 18-DEC-2000 (first entry)

DE 55kd i-antigen nucleotide sequence.

XX Immobilisation antigen; i-antigen; Ichthyophthiriasis; vaccine; ds;

KW white spot disease; freshwater fish; immune response; infection control.

XX Ichthyophthirius multifiliis.

XX WO200046373-A1.

XX 10-AUG-2000.

PF 04-FEB-2000; 2000WO-US02962.

XX 04-FEB-1999; 99US-0118634.

PR 02-MAR-1999; 99US-0122372.

PR 17-MAR-1999; 99US-0124905.

PR 27-APR-1999; 99US-0131121.

XX (UWGE-) UNIV GEORGIA RES FOUND INC.

PA (CORR.) CORNELL RES FOUND INC.

PA (CLAR.) CLARK T G.

PA (DICK.) DICKERSON H W.

PA (LINT.) LINT T.

XX Clark TG, Dickerson HW, Lin T;

PI WPI; 2000-506071/45.

XX Novel i-antigen polypeptides and polynucleotides from Ichthyophthirius

PT multifiliis, useful for prophylaxis and treatment of Ichthyophthirius

PT infection in fish

XX Claim 5; Figure 3; 14pp; English.

XX This invention relates to novel i-antigen polypeptide sequences.

XX I-antigens or immobilisation antigens are common to a variety of

XX hymenostomatid ciliates and their expression varies in response to

XX environmental stimuli. This invention relates to i-antigens in

XX Ichthyophthirius multifiliis, a protozoan which is an obligate parasite

XX of freshwater fish causing ichthyophthiriasis or white spot disease. The

XX invention includes two polypeptide and polynucleotide sequences for two

XX i-antigens, of 48 and 55 kd. Also included in the invention are

XX antibodies capable of binding to the nucleotide sequences and a method

XX for identifying i. multifiliis serotypes using the nucleotide sequences.

Query Match 84.0%; Score 16.8; DB 21; Length 1404;  
Best Local Similarity 90.0%; Pred. No. 5.5e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AATGTTGCTTAAGCTTTT 20  
| | | | | | | | | |  
Db 661 AATGTTGCTTAAGCTTTT 680

RESULT 5  
AAA52136  
ID AAA52136 standard; DNA; 1404 BP.

AC AAA52136;

DT 04-DEC-2000 (first entry)

DE 55 kDa i-antigen gene.

XX BTU; beta-tubulin; protein expression system; negative selection;

KW pacitaxel sensitivity; cell surface; antigen; protozoa; ciliate;

KW live vaccine; Ichthyophthirius multifiliis; immobilization-antigen;

KW i-antigen; freshwater; fish; protozoacide; ds.

XX Ichthyophthirius multifiliis.

XX WO200046381-A1.

XX 10-AUG-2000.

PF 04-FEB-2000; 2000WO-US02966.

XX 04-FEB-1999; 99US-0118634.

PR 02-MAR-1999; 99US-0122372.

PR 17-MAR-1999; 99US-0124905.

PR 27-APR-1999; 99US-0131121.

XX (UWGE-) UNIV GEORGIA RES FOUND INC.

PA (GAER.) GAERTIG J.

PA (DICK.) DICKERSON H W.

PA (CLAR.) CLARK T G.

XX Gaertig J, Dickerson HW, Clark TG;

PI WPI; 2000-514962/46.

XX P-PSDB; AAY971177.

XX Recombinant expression systems for expressing heterologous nucleic

PT acids and producing recombinant protein, comprises nonpathogenic

PT protozoa such as Tetrahymena resistant to paclitaxel

XX Disclosure; Fig 3B; 83pp; English.

XX Tetrahymena thermophila expresses two major beta-tubulin genes (BTU1 and

XX BTU2) which encode identical beta-tubulin proteins. Either of these two

XX genes (but not both at once) can be disrupted without a detectable change

XX in the cell phenotype. A K350L substitution in the BTU1 beta-tubulin

XX protein confers increased resistance to microtubule-depolymerizing drugs

XX and increased sensitivity to paclitaxel, a microtubule-stabilizing drug.

XX Cells carrying the BTU1-K350M allele can be transformed to paclitaxel

XX resistance by gene replacement of BTU1-K350M with a wild-type BTU1 gene

XX fragment, eliminating the need to incorporate a means for positive

XX selection. Where the host organism is not a T. thermophila mutant

XX containing the BTU1-K350M allele, BTU1::neol construct, which

XX substitutes the coding region of the neol gene (conferring resistance to

CC paromycin) for that of BTUL, can be used to generate BTUL gene knockouts  
 CC and for positive selection. Heterologous nucleic acids (especially  
 CC encoding antigenic polypeptides) can be inserted into a BTU gene for  
 CC successful cell-surface expression that is maintained by way of negative  
 CC selection. Preferred expression vectors disrupt the BTUL-1K350M gene by  
 CC homologous recombination-mediated insertion of a heterologous nucleic  
 CC acid, thereby restoring resistance to pacitaxel in the resulting  
 CC transgenic host. Transgenic ciliated protozoa are useful as live vaccines  
 CC for stimulating an immune response in a vertebrate. The transgenic  
 CC protozoan host cells are also useful for producing polyclonal antibodies  
 CC (claimed). In particular, Tetrahymena expressing Ichthyophthirius  
 CC multifiliis immunization antigen (i-antigen) protein on their surface  
 CC are effective vehicles for vaccination of freshwater fish against  
 CC infection by I. multifiliis.

SQ Sequence 1404 BP; 447 A; 241 C; 256 G; 460 T; 0 other;

Query Match Best Local Similarity 84.0%; Score 16.8; DB 21; Length 1404;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 20  
 DB 661 AATGTTGCTTAAGCTTTT 680

RESULT 6  
 AAA97060

ID AAA97060 standard; DNA; 1410 BP.

XX AAA97060;

DT 18-DEC-2000 (first entry)

DE 55KD i-antigen coding region.

XX Immunobilisation antigen; i-antigen; Ichthyophthiriasis; vaccine; ds;

KW white spot disease; freshwater fish; immune response; infection control.

OS Ichthyophthirius multifiliis.

PN WO200046373-A1.

PD 10-AUG-2000.

PF 04-FEB-2000; 2000WO-US02962.

PR 04-FEB-1999; 99US-0118634.

PR 02-MAR-1999; 99US-0122372.

PR 17-MAR-1999; 99US-0124905.

PR 27-APR-1999; 99US-0131121.

PA (UYGE-) UNIV GEORGIA RES FOUND INC.

PA (CORR) CORNELL RES FOUND INC.

PA (DICK/) CLARK T G.

PA (LINT/) DICKERSON H W.

PI CLARK TG, Dickerson HW, Lin T;

DR WPI; 2000-506071/45.

PT Novel i-antigen polypeptides and polynucleotides from Ichthyophthirius

PS infection in fish -

XX Disclosure; Figure 2; 144pp; English.

CC This invention relates to novel i-antigen polypeptide sequences.

CC i-antigens or immunobilisation antigens are common to a variety of

CC environmental stimuli. This invention relates to i-antigens in

CC Ichthyophthirius multifiliis, a protozoan which is an obligate parasite

CC of freshwater fish causing ichthyophthiriasis or white spot disease. The  
 CC invention includes two polypeptide and polynucleotide sequences for two  
 CC i-antigens, of 48 and 55 KD. Also included in the invention are  
 CC antibodies capable of binding to the nucleotide sequences and a method  
 CC for identifying I. multifiliis serotypes using the nucleotide sequences.  
 CC A composition (containing the i-antigen nucleotide) capable of eliciting  
 CC an immune response in fish is useful for prophylaxis, treatment or for  
 CC controlling I. multifiliis infection in fish. Polynucleotide or protein  
 CC vaccines comprising a portion of the amplified product encoding an  
 CC antigenic i-antigen polypeptide obtained is also useful for treating or  
 CC preventing I. multifiliis infection in fish. Sequences AAA97036-A97042,  
 CC and AAA97060, AAA97065 and AAA97089 represent i-antigen genes and gene  
 CC fragments identified in the invention. Sequences AAA97043-A97064  
 CC (excluding AAA97060) and AAA97071-A97088 represent primers used in the  
 CC isolation of the i-antigen gene sequences. Sequences AAB25859-B25889 and  
 CC AAB25893-B25906 represent i-antigen protein and peptide sequences.

SQ Sequence 1410 BP; 449 A; 240 C; 259 G; 462 T; 0 other;

Query Match Best Local Similarity 84.0%; Score 16.8; DB 21; Length 1410;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 20  
 DB 661 AATGTTGCTTAAGCTTTT 680

RESULT 7  
 ABL14642/c

ID ABL14642 standard; cDNA; 4839 BP.

XX ABL14642;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 38408.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-6556860/75.

DR P-PSDB; ABB70539.

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PS genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -

XX Claim 1; SEQ ID NO 38408; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (AAB57757-ABB72072) and the encoded proteins



The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 4839 BP; 1360 A; 1051 C; 1092 G; 1336 T; 0 other;

Query Match 84.0%; Score 16.8; DB 23; Length 4839;  
Best Local Similarity 90.0%; Pred. No. 5.5e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AATGTTGCTTAAGCTTTT 20  
2454 AATGTTGCTTAAGCTTTT 2435

RESULT 8  
ABT10146/c  
ID ABT10146 standard; cDNA; 140167 BP.

ABT10146;

04-DEC-2002 (first entry)

Human breast cancer associated coding sequence SEQ ID NO: 280.

Human; breast specific gene; breast cancer; differential expression;

Cyclostatic; gene therapy; gene; ss.

Homo sapiens.

WO200259271-A2.

01-AUG-2002.

25-JAN-2002; 2002WO-US02176.

25-JAN-2001; 2001US-263757P.

25-APR-2001; 2001US-286090P.

23-MAY-2001; 2001US-292517P.

(GENE-) GENE LOGIC INC.

Orr MS, Nation M, Diggans JC, Zeng W;

WPI; 2002-674803/72.

Diagnosing breast cancer in a patient comprises detecting the level of gene expression in cell or tissue samples, where a differential gene expression is indicative of breast cancer.

Claim 1; SEQ ID NO 280; 260pp + Sequence Listing; English.

The present invention relates to methods of diagnosing breast cancer in a patient, which comprise detecting the level of expression in a tissue sample of two or more genes selected from those shown in ABT09867-ABT1112, where a differential expression of the genes indicates breast cancer. The methods are useful in diagnosing, treating, detecting the progression, and in monitoring treatment of breast cancer in patients. The methods are also useful as a screening tool for agents that modulate the onset or progression of breast cancer. The breast cancer genes may be used as diagnostic markers for the prediction or identification of the malignant state of breast tissue, for confirming the type and progression of cancer, and for drug screening and assays. The present sequence is a coding sequence of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 140167 BP; 45038 A; 27502 C; 26800 G; 40827 T; 0 other;

Query Match 84.0%; Score 16.8; DB 24; Length 140167;  
Best Local Similarity 90.0%; Pred. No. 5.2e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AATGTTGCTTAAGCTTTT 20  
79855 AATGTTGCTTAAGCTTTT 79836

RESULT 9  
ABO88198  
ID ABO88198 standard; cDNA; 154902 BP.

ABO88198;

18-SEP-2002 (first entry)

Human osteoblast differentiation related cDNA SEQ ID NO 105.

Human; osteoblast; stem cell differentiation; bone tissue deposition;  
osteoporosis; osteopathia; ss.

Homo sapiens.

WO200250301-A2.

27-JUN-2002.

18-DEC-2001; 2001WO-US48276.

18-DEC-2000; 2000US-255882P.

24-APR-2001; 2001US-285691P.

(GENE-) GENE LOGIC INC.

(PROC) PROCTER & GAMBLE CO.

Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;

WPI; 2002-557663/59.

Use of genes and their expression profiles associated with osteoblast differentiation for screening modulators bone formation, for diagnosing or treating e.g. osteoporosis, or as markers for the differentiation process.

Claim 1; SEQ ID NO 105; 78pp + Sequence Listing; English.

The invention relates to genes and their expression profiles are used for:

(a) screening modulators of precursor stem cell differentiation into osteoblasts, or bone tissue deposition;

(b) diagnosing abnormal deposition of bone tissue, abnormal rate of osteoblast formation or osteoporosis; or

(c) treating or monitoring treatment of the conditions cited in (b), or monitoring the progression of bone tissue deposition.

Specific conditions include postmenopausal osteoporosis, glucocorticoid osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-induced abnormalities in bone formation or bone loss, conditions that involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis), skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome or fibrous dysplasia. The present sequence is that of an osteoblast differentiation associated cDNA marker of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 154902 BP; 43917 A; 31458 C; 32848 G; 46679 T; 0 other;

Query Match 84.0%; Score 16.8; DB 24; Length 154902;  
Best Local Similarity 90.0%; Pred. No. 5.2e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AATGTTGCTTAAGCTTTT 20  
96683 AATGTTGCTTAAGCTTTT 96702

RESULT 10  
ABK83460/c  
ID ABK83460 standard; cDNA; 165199 BP.  
XX  
AC ABK83460;  
XX  
DT 14-AUG-2002 (first entry)  
XX  
DE Human cDNA differentially expressed in granulocytic cells #31.  
XX  
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;  
KW viral infection; parasitic infection; protozoal infection;  
KW fungal infection; sterile inflammatory disease; psoriasis;  
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;  
KW adult respiratory distress syndrome; inflammatory bowel disease;  
KW Crohn's disease; ulcerative colitis; periodontal disease;  
KW granulocyte activation; chronic inflammation; allergy.  
XX  
OS Homo sapiens.  
XX  
PN WO200228999-A2.  
XX  
PD 11-APR-2002.  
XX  
PF 03-OCT-2001; 2001WO-US30821.  
XX  
PR 03-OCT-2000; 2000US-237189P.  
XX  
PA (GENE-) GENE LOGIC INC.  
XX  
PI Beaser-Barclay Y, Weissman SM, Yamaga S, Vockley J;  
XX  
DR WPI; 2002-435328/46.  
XX  
PT Detecting granulocyte activation by detecting differential expression  
PT of genes associated with granulocyte activation, which serves as  
PT diagnostic markers that is useful for monitoring disease states and  
PT drug toxicity -  
XX  
PS Claim 1; SEQ ID No 31; 114pp; English.

CC parasitic infection, protozoal infection, fungal infection and M5 is  
CC useful for treating one of the above conditions. The present  
CC sequence represents a gene differentially expressed in granulocytes.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 165199 BP; 48510 A; 33223 C; 34406 G; 49060 T; 0 other;  
XX  
Query Match 84.0%; Score 16.8; DB 24; Length 165199;  
Best Local Similarity 90.0%; Pred. No. 5.2e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
XX  
DY 1 AATGTCCTTAACCTTTT 20  
DB 93667 AATGTTCTTACGCAATTTT 93648  
XX  
RESULT 11  
AAS60064  
ID AAS60064 standard; cDNA; 333 BP.  
XX  
AC AAS60064;  
XX  
DT 29-JAN-2002 (first entry)  
XX  
DE Human cancer agent-sensitive marker #65.  
XX  
KW Human; cancer cell marker; TAXOL; cytostatic; tumour; carcinoma;  
KW squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;  
KW lymphocytic leukaemia; lymphoma; plasmacytoma; reticulum cell sarcoma;  
KW Hodgkin's disease; glioma; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200179556-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 13-APR-2001; 2001WO-US12132.  
XX  
PR 14-APR-2000; 2000US-197538P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Lillie J, Brown JL, Bolt A, Van Huffel C;  
XX  
DR WPI; 2001-602933/68.  
XX  
PT Novel nucleic acid, used as a marker to determine the effectiveness of  
PT using TAXOL to treat cancer cell growth in individuals -  
XX  
PS Claim 1; Page 107; 527pp; English.

The invention relates to 1046 novel nucleic acids which are used as  
CC markers for determining the sensitivity of a cancer cell to the  
CC anticancer agent TAXOL. Cancer cells can be treated with TAXOL when  
CC they are shown to express one of the 242 sensitivity markers or the  
CC cells are shown not to express one of the 804 resistance markers.  
CC The methods can be used to determine the effectiveness of TAXOL  
CC in the treatment of cancer cell growth in an individual. The markers  
CC can be used as targets in developing anti-cancer agents such as  
CC chemotherapeutic compounds. The markers can also be used as targets in  
CC developing resistance to agents and exhibit expression of the markers which  
CC anticancer agents developed by the novel method can be used to treat  
CC cancer. Probes based on the markers can be used to detect transcripts or  
CC genomic sequences corresponding to the markers, in the identification of  
CC cells or tissues which mis-express the protein. Cancers which may  
CC be targeted include carcinoma (e.g. squamous cell carcinoma),  
CC sarcoma (e.g. fibrosarcoma) leukaemia (e.g. lymphocytic leukaemia),  
CC lymphoma, plasmacytoma, reticulum cell sarcoma, Hodgkin's disease and

CC tumours (e.g. glioma). The present sequence is one of the 1046  
CC novel cancer cell markers.  
XX  
SQ Sequence 333 BP; 79 A; 63 C; 74 G; 93 T; 24 other;  
Query Match 82.0%; Score 16.4; DB 22; Length 333;  
Best Local Similarity 85.0%; Pred. No. 8.3e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
1 AATGTTGCTTACCTTTT 20  
312 AATTTGCTTGGCTTTT 331  
Db  
RESULT 12  
AAF93687  
ID AAF93687 standard; cDNA; 564 BP.  
AC AAF93687;  
XX  
DT 21-MAY-2001 (first entry)  
XX  
DE cDNA encoding SRT protein isolated from MCF-7 cells SEQ ID 508.  
XX  
KW Human; SRT; gene therapy; gene mapping; tissue typing; ss.  
XX  
OS Homo sapiens.  
XX  
PN W0200107611-A2.  
XX  
PD 01-FEB-2001.  
XX  
PR 21-JUN-2000; 2000WO-US20006.  
XX  
PR 26-JUL-1999; 99US-0145701.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Baker KP, Goddard A, Wood WT;  
XX  
PI WPI; 2001-112729/12.  
XX  
XX New isolated nucleic acid molecule encoding a SRT polypeptide is useful  
XX for production of recombinant SRT polypeptides, gene mapping,  
XX diagnosing genetic disorders and for gene therapy -  
XX  
XX Claim 2; Fig 508; 663bp; English.  
XX  
XX Sequences AAF93180 - AAF93743 represent polynucleotide sequences encoding  
XX human SRT proteins. The cDNA sequences are isolated from various  
XX different human tissue cDNA libraries. The invention relates to a method  
XX for detecting cDNA encoding an SRT protein, a vector containing cDNA  
XX encoding SRT, a host cell transformed with the vector, an isolated SRT  
XX polypeptide, and an antibody which binds to SRT. The polynucleotide  
XX sequence can be used in gene therapy and is useful in the recombinant  
XX production of SRT polypeptides, as a hybridisation probe to screen  
XX libraries to isolate cDNAs with sequence identity to SRT polypeptides, to  
XX map the gene encoding the SRT polypeptides and analysing genetic  
XX disorders, tissue typing and disease tissue detection. The SRT  
XX polynucleotide sequences can be used in polymerase chain reaction,  
XX screening for new therapeutic molecules and generation of antisense RNA  
XX and DNA.  
XX  
SQ Sequence 564 BP; 143 A; 121 C; 126 G; 170 T; 4 other;  
Query Match 82.0%; Score 16.4; DB 22; Length 564;  
Best Local Similarity 89.5%; Pred. No. 8.2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
2 ATGTTGCTTACCTTTT 20  
128 ATGTTGCTTACCTTTT 146  
Db

RESULT 13  
AAC54576  
ID AAC54576 standard; DNA; 1404 BP.  
XX  
AC AAC54576;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 78358.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
XX protein identification; signal transduction pathway;  
XX metabolic pathway; promoter; termination sequence; ss.  
XX  
XX Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
PD 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0128845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX 28-APR-1999; 99US-0130891.  
XX 30-APR-1999; 99US-0131449.  
XX 30-APR-1999; 99US-0132407.  
XX 04-MAY-1999; 99US-0132484.  
XX 05-MAY-1999; 99US-0132485.  
XX 06-MAY-1999; 99US-0132486.  
XX 06-MAY-1999; 99US-0132487.  
XX 07-MAY-1999; 99US-0132863.  
XX 11-MAY-1999; 99US-0134256.  
XX 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134219.  
XX 14-MAY-1999; 99US-0134221.  
XX 14-MAY-1999; 99US-0134370.  
XX 18-MAY-1999; 99US-0134768.  
XX 19-MAY-1999; 99US-0134941.  
XX 20-MAY-1999; 99US-0135124.  
XX 21-MAY-1999; 99US-0135353.  
XX 24-MAY-1999; 99US-0135629.  
XX 25-MAY-1999; 99US-0136021.  
XX 27-MAY-1999; 99US-0136392.  
XX 28-MAY-1999; 99US-0136782.  
XX 01-JUN-1999; 99US-0137222.  
XX 03-JUN-1999; 99US-0137528.  
XX 04-JUN-1999; 99US-0137502.  
XX 07-JUN-1999; 99US-0137724.  
XX 08-JUN-1999; 99US-0138094.  
XX 10-JUN-1999; 99US-0138540.  
XX 10-JUN-1999; 99US-0138847.  
XX 14-JUN-1999; 99US-0139119.  
XX 16-JUN-1999; 99US-0139452.  
XX 16-JUN-1999; 99US-0139453.  
XX 17-JUN-1999; 99US-0139492.  
XX 18-JUN-1999; 99US-0139454.  
XX 18-JUN-1999; 99US-0139455.  
XX 18-JUN-1999; 99US-0139456.  
XX 18-JUN-1999; 99US-0139457.

```
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142955.
PR 08-JUL-1999; 99US-0142990.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145195.
PR 26-JUL-1999; 99US-0145218.
PR 27-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 28-JUL-1999; 99US-0145813.
PR 02-AUG-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154019.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156599.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159325.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 16-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 82.0%; Score 16.4; DB 21; Length 1404;
Best Local Similarity 94.4%; Pred. No. 8.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TGTGCTTAAGCTTTT 20
Db 520 TGTGCTTAAGCTTTT 537

RESULT 14
ABL17270
ID ABL17270 standard; DNA; 2613 BP.
XX AC ABL17270;
XX
```

DT 26-MAR-2002 (first entry)  
 XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 3283.  
 DE Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ds.  
 XX Drosophila melanogaster.  
 OS WO200171042-A2.  
 XX  
 XX  
 PD 27-SEP-2001.  
 XX  
 XX 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX WPI; 2001-656860/75.  
 DR  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Claim 1; SEQ ID NO 3283; 21bp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins  
 CC (AB857737-AB872072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SQ Sequence 2613 BP; 874 A; 462 C; 460 G; 817 T; 0 other;  
 Query Match 82.0%; Score 16.4; DB 23; Length 2613;  
 Best Local Similarity 94.4%; Pred. No. 8.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AATGTCCTTAAGCTTTT 18  
 DB 2077 AATGTCCTTAAGCTTTT 2094  
 RESULT 15  
 ABL22326  
 ID ABL22326 standard; DNA; 3945 BP.  
 AC ABL22326;  
 XX  
 XX 26-MAR-2002 (first entry)  
 DT  
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 18451.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ds.  
 XX Drosophila melanogaster.  
 OS WO200171042-A2.  
 XX  
 XX  
 PD 27-SEP-2001.  
 XX  
 XX 23-MAR-2001; 2001WO-US09231.  
 PF

XX 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 XX (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX WPI; 2001-656860/75.  
 DR  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Claim 1; SEQ ID NO 18451; 21bp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins  
 CC (AB857737-AB872072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SQ Sequence 3945 BP; 1117 A; 835 C; 876 G; 1117 T; 0 other;  
 Query Match 82.0%; Score 16.4; DB 23; Length 3945;  
 Best Local Similarity 94.4%; Pred. No. 8.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 TGTTCCTTAAGCTTTT 20  
 DB 2533 TGTTCCTTAAGCTTTT 2550

Search completed: August 1, 2003, 13:55:21  
 Job time : 10.6987 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:05:17 ; Search time 5.18607 Seconds  
(without alignments)  
7955.924 Million cell updates/sec

Title: US-09-834-291-14  
Perfect score: 20  
Sequence: 1 aatgtgcttaagctttttc 20

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database: Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_NEW\_PUB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	US-09-834-291-14	Sequence 14, Appl
2	20	100.0	40	US-09-834-291-20	Sequence 20, Appl
3	20	100.0	40	US-09-834-291-28	Sequence 28, Appl
4	20	100.0	2827	US-09-834-291-4	Sequence 4, Appl
5	20	100.0	3212	US-09-834-291-1	Sequence 1, Appl
6	17.4	87.0	474	US-10-027-632-66459	Sequence 36459, A
7	17.4	84.0	591	US-10-027-632-69220	Sequence 69220, A
8	16.8	84.0	20	US-09-834-291-15	Sequence 15, Appl
9	16.8	84.0	553	US-10-027-632-72944	Sequence 72944, A
10	16.8	84.0	960	US-10-027-632-9152	Sequence 9152, Ap
11	16.8	84.0	1055	US-10-027-632-250885	Sequence 250885,
12	16.8	84.0	1055	US-10-027-632-250886	Sequence 250886,
13	16.4	82.0	333	US-09-834-975-65	Sequence 65, Appl
14	16.4	82.0	549	US-10-066-543-2878	Sequence 2878, Ap
15	16.4	82.0	602	US-10-027-632-135561	Sequence 135561,
16	16.4	82.0	602	US-10-027-632-135562	Sequence 135562,

17	16.4	82.0	785	13	US-10-027-632-157839	Sequence 157839,
18	16.4	82.0	118067	15	US-10-081-327-32	Sequence 32, Appl
19	16.4	82.0	1503841	9	US-09-795-668-1	Sequence 1, Appl
20	16.4	82.0	1503841	9	US-09-795-668-1	Sequence 1, Appl
21	16.4	82.0	1503841	10	US-09-946-807-1	Sequence 1, Appl
22	16.4	82.0	1501042	13	US-10-027-632-58064	Sequence 58064, A
23	16	80.0	652	13	US-10-027-632-278366	Sequence 278366,
24	16	80.0	6165	10	US-09-880-107-3866	Sequence 3866, Ap
25	16	80.0	83450	15	US-09-811-469-3	Sequence 3, Appl
26	16	80.0	161652	15	US-10-081-327-40	Sequence 40, Appl
27	15.8	79.0	437	13	US-10-027-632-267598	Sequence 267598,
28	15.8	79.0	441	13	US-10-027-632-283756	Sequence 283756,
29	15.8	79.0	468	10	US-09-864-824A-148	Sequence 148, App
30	15.8	79.0	468	10	US-09-864-824A-148	Sequence 148, App
31	15.8	79.0	475	13	US-10-027-632-27810	Sequence 27810, Ap
32	15.8	79.0	475	13	US-10-027-632-27810	Sequence 27810, Ap
33	15.8	79.0	559	13	US-10-027-632-200450	Sequence 200450,
34	15.8	79.0	559	13	US-10-027-632-200451	Sequence 200451,
35	15.8	79.0	589	13	US-10-027-632-290172	Sequence 290172,
36	15.8	79.0	592	13	US-10-027-632-107268	Sequence 107268,
37	15.8	79.0	625	13	US-10-027-632-211910	Sequence 211910,
38	15.8	79.0	643	13	US-10-027-632-211911	Sequence 211911,
39	15.8	79.0	643	13	US-10-027-632-233945	Sequence 233945,
40	15.8	79.0	643	13	US-10-027-632-27458	Sequence 27458, A
41	15.8	79.0	663	13	US-10-027-632-237201	Sequence 237201,
42	15.8	79.0	663	13	US-10-027-632-237202	Sequence 237202,
43	15.8	79.0	667	13	US-10-027-632-237203	Sequence 237203,
44	15.8	79.0	677	13	US-10-027-632-290173	Sequence 290173,
45	15.8	79.0	677	13	US-10-027-632-32623	Sequence 32623, A

## ALIGNMENTS

RESULT 1  
US-09-834-291-14  
Sequence 14, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834, 291  
PRIOR FILING DATE: 2001-08-21  
CURRENT FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 14  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-14

Query Match 100.0%; Score 20; DB 9; Length 20;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTGCTTAAGCTTTT 20  
DB 1 AATGTGCTTAAGCTTTT 20

RESULT 2  
US-09-834-291-20  
Sequence 20, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter

APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: P53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 20  
LENGTH: 40  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-20

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 40;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 20  
DB 11 AATGTTGCTTAAGCTTTT 30

RESULT 3  
US-09-834-291-28  
Sequence 28, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Kramer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: P53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 28  
LENGTH: 40  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-28

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 40;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 20  
DB 11 AATGTTGCTTAAGCTTTT 30

RESULT 4  
US-09-834-291-4  
Sequence 4, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Kramer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: P53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343

PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2827  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-4

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 2827;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 20  
DB 437 AATGTTGCTTAAGCTTTT 456

RESULT 5  
US-09-834-291-1  
Sequence 1, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Kramer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: P53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 3212  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-1

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 3212;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 20  
DB 437 AATGTTGCTTAAGCTTTT 456

RESULT 6  
US-10-027-632-36459/c  
Sequence 36459, Application US/10027632  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
Polymorphisms in the Human Genome  
FILE REFERENCE: 106827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/216,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23



PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 36459  
LENGTH: 474  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-36459

Query Match 87.0%; Score 17.4; DB 13; Length 474;  
Best Local Similarity 94.7%; Pred. No. 3.3e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ATGTTGCTTAAGCTTTT 20  
DB 337 ATGTTGCTTAAGCTTTT 319

RESULT 7  
US-10-027-632-69220/C  
Sequence 69220, Application US/10027632

GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
TITLE OF INVENTION: Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT FILING DATE: 2002-04-30  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 69220  
LENGTH: 591  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-69220

Query Match 87.0%; Score 17.4; DB 13; Length 591;  
Best Local Similarity 94.7%; Pred. No. 3.4e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ATGTTGCTTAAGCTTTT 20  
DB 454 ATGTTGCTTAAGCTTTT 436

RESULT 8  
US-09-834-291-15  
Sequence 15, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291

CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: ECT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 15  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-15

Query Match 84.0%; Score 16.8; DB 9; Length 20;  
Best Local Similarity 90.0%; Pred. No. 3.9e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGCTTTT 20  
DB 1 AATGTTCTTAAGCTTTT 20

RESULT 9  
US-10-027-632-72944/C  
Sequence 72944, Application US/10027632

GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
TITLE OF INVENTION: Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT FILING DATE: 2002-04-30  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 72944  
LENGTH: 553  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)-(553)  
OTHER INFORMATION: n = A,T,C or G  
US-10-027-632-72944

Query Match 84.0%; Score 16.8; DB 13; Length 553;  
Best Local Similarity 90.0%; Pred. No. 6e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGCTTTT 20  
DB 38 AATGTTCTTAAGCTTTT 19

RESULT 10  
US-10-027-632-9152/C  
Sequence 9152, Application US/10027632  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

```

; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9152
; LENGTH: 980
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-9152
```

```

Query Match      84.0%; Score 16.8; DB 13; Length 980;
Best Local Similarity 90.0%; Pred. No. 6.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

1 AATGTTGCTTAAGCTTTT 20
|||||
111 AATGTTCTGAAGCTTTT 92
```

```

RESULT 11
US-10-027-632-250885/c
; Sequence 250885, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 250885
; LENGTH: 1055
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-250885
```

```

Query Match      84.0%; Score 16.8; DB 13; Length 1055;
Best Local Similarity 90.0%; Pred. No. 6.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AATGTTGCTTAAGCTTTT 20
|||||
```

```

Db      464 AATGTTGCTTAAGCTTTT 445
```

```

RESULT 12
US-10-027-632-250886/c
; Sequence 250886, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 250886
; LENGTH: 1055
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-250886
```

```

Query Match      84.0%; Score 16.8; DB 13; Length 1055;
Best Local Similarity 90.0%; Pred. No. 6.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

1 AATGTTGCTTAAGCTTTT 20
|||||
464 AATGTTGCTTAAGCTTTT 445
```

```

RESULT 13
US-09-834-975-65
; Sequence 65, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Hufel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-016B
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US/09/834,975
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(333)
; OTHER INFORMATION: n = A,T,C or G
; US-09-834-975-65
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 02:21:37 ; Search time 1.65066 Seconds  
(without alignments)  
5347.959 Million cell updates/sec

Title: US-09-834-291-14

Perfect score: 20

Sequence: 1 aatgttgcttaagcttttc 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.8	84.0	1077	4	US-09-328-352-2249
2	16	80.0	83450	4	US-09-811-469-3
3	15.8	79.0	186	4	US-09-328-352-3726
4	15.8	79.0	1485	4	US-09-328-352-1158
5	15.8	79.0	1551	4	US-09-328-352-2348
6	15.8	79.0	5134	2	US-08-310-912A-157
7	15.8	79.0	5134	3	US-09-301-085-157
8	15.8	79.0	5134	5	PCT-US95-04589-157
9	15.8	79.0	5475	2	US-08-680-327-1
10	15.8	79.0	5475	3	US-09-228-246-3
11	15.8	79.0	10968	2	US-08-680-327-2
12	15.8	79.0	10968	3	US-09-228-246-1
13	15.8	79.0	580073	4	US-08-545-528D-1
14	15.8	79.0	1664976	4	US-08-916-421B-1
15	15.4	77.0	294	4	US-09-016-434-1038
16	15.4	77.0	3024	4	US-09-620-312D-534
17	15.2	76.0	521	4	US-09-643-597-214
18	15.2	76.0	521	4	US-09-480-884A-214
19	15.2	76.0	521	4	US-09-542-615A-214
20	15.2	76.0	521	4	US-09-606-421B-214
21	15.2	76.0	602	3	US-09-040-964-11
22	15.2	76.0	602	3	US-09-123-912-11
23	15.2	76.0	602	4	US-09-643-597-11
24	15.2	76.0	602	4	US-09-480-884A-11
25	15.2	76.0	602	4	US-09-542-615A-11
26	15.2	76.0	602	4	US-09-606-421B-11
27	15.2	76.0	606	3	US-09-040-984-55

28	15.2	76.0	606	4	US-09-123-912-55	Sequence 55, Appl
29	15.2	76.0	606	4	US-09-643-597-55	Sequence 55, Appl
30	15.2	76.0	606	4	US-09-480-884A-55	Sequence 55, Appl
31	15.2	76.0	606	4	US-09-542-615A-55	Sequence 55, Appl
32	15.2	76.0	606	4	US-09-606-421B-55	Sequence 108, App
33	15.2	76.0	645	4	US-09-522-217-108	Sequence 1717, Ap
34	15.2	76.0	822	4	US-09-328-352-1177	Sequence 1, Appl
35	15.2	76.0	1011	3	US-08-686-528A-1	Sequence 1, Appl
36	15.2	76.0	1011	4	US-09-456-287-1	Sequence 10, Appl
37	15.2	76.0	1011	4	US-08-487-429A-10	Sequence 119, App
38	15.2	76.0	1011	5	PCT-US96-05320A-119	Sequence 1995, App
39	15.2	76.0	1026	4	US-09-134-001C-1995	Sequence 101, App
40	15.2	76.0	1026	4	US-09-280-116-101	Sequence 4, Appl
41	15.2	76.0	1856	4	US-08-845-258-4	Sequence 4, Appl
42	15.2	76.0	1991	3	US-08-990-571-4	Sequence 4, Appl
43	15.2	76.0	1991	4	US-08-723-142A-4	Sequence 4, Appl
44	15.2	76.0	1991	4	US-09-528-784A-4	Sequence 4, Appl
45	15.2	76.0	1991	4	US-09-569-098A-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-09-328-352-2249/c  
; Sequence 2249, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 2249  
; LENGTH: 1077  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-2249

Query Match 84.0%; Score 16.8; DB 4; Length 1077;  
Best Local Similarity 90.0%; Pred. No. 65;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 AATGTTGCTTAAGCTTTT 20  
Db 616 AATGTTGCTTAAGCTTTT 597

RESULT 2  
US-09-811-469-3  
; Sequence 3, Application US/09811469  
; Patent No. 6551809  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,  
; FILE REFERENCE: CLO01171  
; CURRENT APPLICATION NUMBER: US/09/811,469  
; CURRENT FILING DATE: 2001-03-20  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3 83450  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(83450)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-811-469-3

Query Match 80.0%; Score 16; DB 4; Length 83450;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AATGTTGCTTAAGCTTT 16  
DB 62943 AATGTTGCTTAAGCTT 62958

RESULT 3  
US-09-328-352-3726/c  
Sequence 3726, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 3726  
LENGTH: 186  
TYPE: DNA  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-3726

Query Match 79.0%; Score 15.8; DB 4; Length 186;  
Best Local Similarity 89.5%; Pred. No. 1.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 AATGTTGCTTAAGCTTTT 20  
DB 100 ATGATGCTTTAGCTTTT 82

RESULT 4  
US-09-328-352-1158  
Sequence 1158, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 1158  
LENGTH: 1485  
TYPE: DNA  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-1158

Query Match 79.0%; Score 15.8; DB 4; Length 1485;  
Best Local Similarity 89.5%; Pred. No. 1.8e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AATGTTGCTTAAGCTTTT 19  
DB 1085 AATGCGGCTTAAGCTTTT 1103

RESULT 5  
US-09-328-352-2348/c  
Sequence 2348, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTC99-03PA

CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 2348  
LENGTH: 1551  
TYPE: DNA  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-2348

Query Match 79.0%; Score 15.8; DB 4; Length 1551;  
Best Local Similarity 89.5%; Pred. No. 1.8e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AATGTTGCTTAAGCTTTT 19  
DB 1492 AAGCTTGTAAAGCTTTT 1474

RESULT 6  
US-08-310-912A-157/c  
Sequence 157, Application US/08310912A  
Patent No. 5981730  
GENERAL INFORMATION:  
APPLICANT: Ausubel, Frederick M.  
APPLICANT: Staskiewicz, Brian J.  
APPLICANT: Brent, Andrew F.  
APPLICANT: Dahlbeck, Douglas  
APPLICANT: Katagiri, Fumitaki  
APPLICANT: Kunkel, Barbara N.  
APPLICANT: Mindrinos, Michael N.  
APPLICANT: Yu, Guo-Liang  
TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND DETECTION  
NUMBER OF SEQUENCES: 208  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2904  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/310,912A  
FILING DATE: September 22, 1994  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/227,360  
FILING DATE: April 13, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lech, Karen F.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/254001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 100254  
INFORMATION FOR SEQ ID NO: 157:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5134 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-310-912A-157

Query Match 79.0%; Score 15.8; DB 2; Length 5134;  
Best Local Similarity 89.5%; Pred. No. 1.8e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 19  
|||||  
Db 462 AATGTTGATAAAGCTTTT 444

US-09-301-085-157/C  
Sequence 157, Application US/09301085

PATENT No. 6262248  
GENERAL INFORMATION:  
APPLICANT: Ausubel, Frederick M.  
APPLICANT: Staekawicz, Brian J.  
APPLICANT: Brent, Andrew F.  
APPLICANT: Dahlbeck, Douglas  
APPLICANT: Katagiri, Fumaki  
APPLICANT: Kunkel, Barbara N.  
APPLICANT: Mindrinos, Michael N.  
APPLICANT: Yu, Guo-Liang  
TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND  
FILE OF INVENTION: DETECTION METHODS  
FILE REFERENCE: 00786/254002  
CURRENT APPLICATION NUMBER: US/09/301,085  
CURRENT FILING DATE: 1999-04-28  
EARLIER APPLICATION NUMBER: 08/310,912  
EARLIER FILING DATE: 1994-09-22  
EARLIER APPLICATION NUMBER: 08/227,360  
EARLIER FILING DATE: 1994-04-13  
NUMBER OF SEQ ID NOS: 208  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 157  
LENGTH: 5134  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-301-085-157

Query Match 79.0%; Score 15.8; DB 3; Length 5134;  
Best Local Similarity 89.5%; Pred. No. 1.8e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 19  
|||||  
Db 462 AATGTTGATAAAGCTTTT 444

## RESULT 8

PCT-US95-04589-157/C  
Sequence 157, Application PC/TUS9504589

GENERAL INFORMATION:  
APPLICANT: Ausubel, Frederick M.  
APPLICANT: Staekawicz, Brian J.  
APPLICANT: Brent, Andrew F.  
APPLICANT: Dahlbeck, Douglas  
APPLICANT: Katagiri, Fumaki  
APPLICANT: Kunkel, Barbara N.  
APPLICANT: Mindrinos, Michael N.  
APPLICANT: Yu, Guo-Liang  
TITLE OF INVENTION: RPS2 GENE AND USES THEREOF  
NUMBER OF SEQUENCES: 201  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street Suite 3100  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2904  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04589

FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,360  
FILING DATE: 13-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/230001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 100254

INFORMATION FOR SEQ ID NO: 157:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5134 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
PCT-US95-04589-157

Query Match 79.0%; Score 15.8; DB 5; Length 5134;  
Best Local Similarity 89.5%; Pred. No. 1.8e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 19  
|||||  
Db 462 AATGTTGATAAAGCTTTT 444

US-08-680-327-1/C  
Sequence 1, Application US/08680327

PATENT No. 5859321  
GENERAL INFORMATION:  
APPLICANT: Staekawicz, Brian S., Oldroyd, Giles Edward,  
APPLICANT: Salmeron, John M., Rommens, Catus  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PLANT  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klarkist Sparkman Campbell Leigh &  
ADDRESS: Winston  
STREET: One World Trade Center  
STREET: 121 S.W. Salmon Street  
STREET: Suite 1600  
CITY: Portland  
STATE: Oregon  
COUNTRY: United States of America  
ZIP: 97204

COMPUTER READABLE FORM:  
MEDIUM TYPE: Disk, 3-1/2 inch  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/680,327  
FILING DATE: July 11, 1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/310,912  
FILING DATE: September 22, 1994  
CLASSIFICATION: 800  
APPLICATION NUMBER: 08/227,360  
FILING DATE: April 13, 1994  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Alan E.  
REGISTRATION NUMBER: 35,123  
REFERENCE/DOCKET NUMBER: 5151-45038  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (503) 226-7391

TELEFAX: (503) 228-9446  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5475 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double stranded  
TOPOLOGY: linear  
US-08-680-327-1

Query Match  
Best Local Similarity 79.0%; Score 15.8; DB 2; Length 5475;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 19  
DB 2508 AATGTTGATTAAGCTTTT 2490

RESULT 10  
US-09-228-246-3/c  
Sequence 3, Application US/09228246  
Patent No. 6245510  
GENERAL INFORMATION:  
APPLICANT: Staskawicz, B. S. et al.  
TITLE OF INVENTION: PRF Protein and Nucleic Acid Sequences: Compositions  
FILE REFERENCE: 51700  
CURRENT APPLICATION NUMBER: US/09/228,246  
EARLIER FILING DATE: 1999-01-11  
EARLIER APPLICATION NUMBER: 08/680,327  
EARLIER FILING DATE: 1996-07-11  
EARLIER APPLICATION NUMBER: 08/310,912  
EARLIER FILING DATE: 1994-09-22  
EARLIER APPLICATION NUMBER: 08/227,360  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 5475  
TYPE: DNA  
ORGANISM: Lycopersicon esculentum  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1) .. (5475)  
US-09-228-246-3

Query Match  
Best Local Similarity 79.0%; Score 15.8; DB 3; Length 5475;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 19  
DB 2508 AATGTTGATTAAGCTTTT 2490

RESULT 11  
US-08-680-327-2/c  
Sequence 2, Application US/08680327  
Patent No. 5859321  
GENERAL INFORMATION:  
APPLICANT: Staskawicz, Brian S., Oldroyd, Giles Edward,  
APPLICANT: Salmeron, John M., Rommens, Calus  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PLANT  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klarkquist Sparkman Campbell Leigh &  
STREET: One World Trade Center  
STREET: 121 S.W. Salmon Street  
CITY: Suite 1600  
CITY: Portland  
STATE: Oregon

COUNTRY: United States of America  
ZIP: 97204  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Disk, 3-1/2 inch  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/680,327  
FILING DATE: July 11, 1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/310,912  
FILING DATE: September 22, 1994  
CLASSIFICATION: 800  
APPLICATION NUMBER: 08/227,360  
FILING DATE: April 13, 1994  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Alan. E.  
REGISTRATION NUMBER: 35,123  
REFERENCE/DOCKET NUMBER: 5151-45038  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (503) 226-7391  
TELEFAX: (503) 228-9446  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10968 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double stranded  
TOPOLOGY: linear  
US-08-680-327-2

Query Match  
Best Local Similarity 79.0%; Score 15.8; DB 2; Length 10968;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTGCTTAAGCTTTT 19  
DB 6386 AATGTTGATTAAGCTTTT 6368

RESULT 12  
US-09-228-246-1/c  
Sequence 1, Application US/09228246  
Patent No. 6245510  
GENERAL INFORMATION:  
APPLICANT: Staskawicz, B. S. et al.  
TITLE OF INVENTION: PRF Protein and Nucleic Acid Sequences: Compositions  
FILE REFERENCE: 51700  
CURRENT APPLICATION NUMBER: US/09/228,246  
EARLIER FILING DATE: 1999-01-11  
EARLIER APPLICATION NUMBER: 08/680,327  
EARLIER FILING DATE: 1996-07-11  
EARLIER APPLICATION NUMBER: 08/310,912  
EARLIER FILING DATE: 1994-09-22  
EARLIER APPLICATION NUMBER: 08/227,360  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 10968  
TYPE: DNA  
ORGANISM: Lycopersicon esculentum  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (3879) .. (8186)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (8300) .. (9466)  
US-09-228-246-1



```
Query Match      79.0%; Score 15.8; DB 3;
Best Local Similarity 89.5%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 AATGTTGCTTAAGCTTTT 19
      |||||
Db      6386 AATGTTGATTAAGCTTTT 6386

RESULT 13
US-08-545-528D-1
; Sequence 1, Application US/08545528D
; Patent No. 6537773
; GENERAL INFORMATION:
; APPLICANT: Fraser et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragment
; Patent No. 6537773
; FILE REFERENCE: Thereof, and Uses Thereof
; CURRENT APPLICATION NUMBER: US/08/545,528D
; CURRENT FILING DATE: 1995-10-19
; PRIOR APPLICATION NUMBER: US 08/488,018
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/473,545
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 580073
; TYPE: DNA
; ORGANISM: Mycoplasma genitalium
US-08-545-528D-1

Query Match      79.0%; Score 15.8; DB 4;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 AATGTTGCTTAAGCTTTT 20
      |||||
Db      2199 AATGTTCTTAAGCTTTT 2217

RESULT 14
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
```

```
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (622708)..(622708)
```

OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (657081)..(657081)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (657203)..(657203)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (674435)..(674435)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (682442)..(682442)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (713652)..(713652)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (741684)..(741684)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (779455)..(779455)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (779676)..(779676)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (855539)..(855539)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (871619)..(871619)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1084830)..(1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1096846)..(1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1119881)..(1119881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1130881)..(1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349473)..(1349473)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349491)..(1349491)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1470091)..(1470091)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1569020)..(1569020)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1602912)..(1602912)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1603734)..(1603734)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1637998)..(1637998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1664854)..(1664854)  
OTHER INFORMATION: n equals a, t, c, or g

US-08-916-421B-1

Query Match

Best Local Similarity 79.0%; Score 15.8; DB 4; Length 1664976;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AATGTCCTTAAGCTTTT 19  
|||||  
Db 789019 AATGTCCTTAAGCTTTT 789001

RESULT 15

US-09-016-434-1038  
Sequence 1038, Application US/09016434  
Patent No. 6500938

GENERAL INFORMATION:

APPLICANT: Janice Au-Young

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

NUMBER OF SEQUENCES: 1490

CORRESPONDENCE ADDRESS:

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/016,434

FILING DATE: HEREWITH

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0002 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1038:

SEQUENCE CHARACTERISTICS:

LENGTH: 294 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: MUSCNOT02

CLONE: 975377

US-09-016-434-1038

Query Match 77.0%; Score 15.4; DB 4; Length 294;  
Best Local Similarity 94.1%; Pred. No. 2.6e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AATGTCCTTAAGCTTTT 18  
|||||  
Db 201 AATGTCCTTAAGCTTTT 217

Search completed: August 1, 2003, 08:37:17  
Job time: 5.65066 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:37:27 ; Search time 78.1573 Seconds  
(without alignments)  
10468.541 Million cell updates/sec

Title: US-09-834-291-15  
Perfect score: 20  
Sequence: 1 aatgttctaagattttt 20

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl : \*

- 1: gb\_ba:\*
- 2: gb\_hcg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vl:\*
- 15: gb\_da:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_or:\*
- 21: em\_ov:\*
- 22: em\_pal:\*
- 23: em\_ph:\*
- 24: em\_pl:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vl:\*
- 30: em\_hcg\_hum:\*
- 31: em\_hcg\_inv:\*
- 32: em\_hcg\_other:\*
- 33: em\_hcg\_mus:\*
- 34: em\_hcg\_pln:\*
- 35: em\_hcg\_rod:\*
- 36: em\_hcg\_mam:\*
- 37: em\_hcg\_vrt:\*
- 38: em\_gy:\*
- 39: em\_hugo\_hum:\*
- 40: em\_hugo\_mus:\*
- 41: em\_hugo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	6 AX026103	AX026103 Sequence
2	18.4	92.0	645	6 AR174645	AR174645 Sequence
3	18.4	92.0	1454	10 RNCND1	X75207 R.novvegicu
4	18.4	92.0	2358	10 BC044841	BC044841 Mus muscu
5	18.4	92.0	3737	6 AX685351	AX685351 Sequence
6	18.4	92.0	3737	10 S78355	S78355 Cyl-1-cycl1
7	18.4	92.0	3751	10 RATCYCD1	D14014 Rattus ratt
8	18.4	92.0	5019	3 AF060234	AF060234 Valrimorp
9	18.4	92.0	5226	3 AC014915	AC014915 Drosophi1
10	18.4	92.0	5854	3 DMHSH333	X81207 D.melanogas
11	18.4	92.0	6164	6 AX344941	AX344941 Sequence
12	18.4	92.0	16243	6 AC014426	AC014426 Drosophi1
13	18.4	92.0	28486	6 AX695350	AX695350 Sequence
14	18.4	92.0	30033	3 CEK10H10	Z8236 Ctenorhadi
15	18.4	92.0	33930	3 CEK04H4	Z27078 Ctenorhadi
16	18.4	92.0	34700	2 AC019929	AC019929 Drosophi1
17	18.4	92.0	80707	2 AC016111	AC016111 Homo sapi
18	18.4	92.0	86576	2 AC112247	AC112247 Homo sapi
19	18.4	92.0	90000	2 AC068803	AC068803 Homo sapi
20	18.4	92.0	100721	5 AL591520	AL591520 Zebrafish
21	18.4	92.0	105922	2 DMHR21N6	AL121842 Drosophi1
22	18.4	92.0	105932	2 AC010050	AC010050 Drosophi1
23	18.4	92.0	124636	2 AC091690	AC091690 Oryza sat
24	18.4	92.0	125981	2 AP003875	AP003875 Oryza sat
25	18.4	92.0	145000	2 AC124784	AC124784 Homo sapi
26	18.4	92.0	144000	5 AL935296	AL935296 Homo sapi
27	18.4	92.0	147042	2 AC121705	AL935296 Zebrafish
28	18.4	92.0	150875	2 AP005726	AP005726 Rattus no
29	18.4	92.0	153037	2 AP005094	AP005094 Oryza sat
30	18.4	92.0	153065	2 AC008380	AC008380 Homo sapi
31	18.4	92.0	153890	2 AC008380	AC008380 Homo sapi
32	18.4	92.0	153903	9 AC008503	AC008503 Homo sapi
33	18.4	92.0	153906	10 AF384675	AF384675 Mus muscu
34	18.4	92.0	154480	2 AP004565	AP004565 Oryza sat
35	18.4	92.0	155053	2 AP005523	AP005523 Oryza sat
36	18.4	92.0	155383	2 AP004396	AP004396 Oryza sat
37	18.4	92.0	158241	2 AC021804	AC021804 Homo sapi
38	18.4	92.0	162626	2 AC138551	AC138551 Danio rer
39	18.4	92.0	163280	2 AC141497	AC141497 Rattus no
40	18.4	92.0	163768	2 AC094692	AC094692 Rattus no
41	18.4	92.0	168299	3 AC113250	AC113250 Drosophi1
42	18.4	92.0	168538	2 AC133565	AC133565 Homo sapi
43	18.4	92.0	168723	2 AC134719	AC134719 Rattus no
44	18.4	92.0	170280	2 AP005520	AP005520 Oryza sat
45	18.4	92.0	170280	2 AP005520	AP005520 Oryza sat

## ALIGNMENTS

RESULT 1  
AX026103  
LOCUS  
Sequence 15 from Patent DE19847779.  
ACCESION  
AX026103  
VERSION  
AX026103.1 GI:10187534  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Mueller-Schilling, M., Krammer, P. and Oren, M.  
Novel receptor dna useful for identifying apoptosis-modulating  
substances potentially useful for cancer chemotherapy

JOURNAL Patent: DE 19847779-C 15 03-FEB-2000;  
DEUTSCHES KREBSFORSCH (DE)

FEATURES  
source  
Location/Qualifiers  
1. 20  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 5 a 1 c 2 g 12 t  
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.e+03;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20  
1 AATGTTCTTAAGATTTT 20

Db 1 AATGTTCTTAAGATTTT 20

RESULT 2  
AR174645/c AR174645 645 bp DNA linear PAT 17-DEC-2001  
LOCUS Sequence 108 from patent US 6307024.  
DEFINITION AR174645  
ACCESSION AR174645  
VERSION AR174645.1 GI:17914965  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
AUTHORS  
1 (bases 1 to 645)  
Novak,J.E., Presnell,S.R., Sprecher,C.A., Foster,D.C., Holly,R.D.,  
Gross,J.A., Johnson,J.V., Nelson,A.J., Dillon,S.R. and  
Hammond,A.K.  
Cytokine zaihal1 Ligand  
Patent: US 6307024-A 108 23-OCT-2001;  
Location/Qualifiers  
1. 645  
/organism="unknown"

BASE COUNT 244 a 107 c 89 g 204 t 1 others  
ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 645;  
Best Local Similarity 95.0%; Pred. No. 2.9e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20  
1 AATGTTCTTAAGATTTT 364

Db 383 AATGTTCTTAAGATTTT 364

RESULT 3  
RNCND1/c RNCND1 1454 bp mRNA linear ROD 17-FEB-1995  
LOCUS R.norvegicus CND1 mRNA for cyclin D1.  
DEFINITION X75207.1 GI:473122  
ACCESSION X75207.1  
VERSION X75207.1  
KEYWORDS Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE  
AUTHORS  
1 Bianchi,S., Fabiani,S., Muratori,M., Arnold,A., Sakaguchi,K.,  
Miki,T. and Brandt,M.L.  
Calcium modulates the cyclin D1 expression in a rat parathyroid  
cell line  
Biochem Biophys. Res. Commun. 204 (2), 691-700 (1994)

JOURNAL MEDLINE  
PUBMED 95071382  
7980531  
2 (bases 1 to 1454)  
Brandt,M.  
TITLE Direct Submission

JOURNAL Submitted (23-SEP-1993) M. Brandt, Endocrine Unit, University of  
Florence, Dept of Clinical Physiopathology, Viale Pieraccini 6,  
50139 Firenze, ITALY

FEATURES  
source  
Location/Qualifiers  
1. 1454  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/cell\_line="patented rat parathyroid cell line"  
1. 1454  
/gene="CND1"  
153. 1040  
/gene="CND1"  
/codon\_start=1  
/product="CYCLIN D1"  
/protein\_id="CA53020.1"  
/db\_xref="GI:473123"  
/db\_xref="SWISS-PROT:P3948"  
/translation="MEHQLCCVEETIRRAYPPTNLINDRVLRAMLKTEETCAPSVY  
FKCVOREIVPSMKRIIVATWMLVEYCEBOKCEEFPIAMNYLDPLSLIPKSRLOIL  
GATCMFVASMKEKTIPLTAEKICTYDNGSIRPELLQWELLVKKWNLAAATPHDF  
IEHFLSKMPEADENKQIIRKHAOTFVALCATDVKFTSNPESVVAAGSVVAAQGLING  
SPNFRFASCTRTTHFLSRVTKCDPCIRACQEQIETALLBSLSLROAQONIDPKATEEGE  
VEREALACTPTDVRVDI"

BASE COUNT 358 a 397 c 394 g 305 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 1454;  
Best Local Similarity 95.0%; Pred. No. 2.4e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20  
1 AATGTTCTTAAGATTTT 1430

Db 1449 AATGTTCTTAAGATTTT 1430

RESULT 4  
BC044841/c BC044841 2358 bp mRNA linear ROD 28-JAN-2003  
LOCUS Mus musculus, cyclin D1, clone MGC:7003 IMAGE:3155470, mRNA,  
DEFINITION complete cds.  
ACCESSION BC044841  
VERSION BC044841.1 GI:27924088  
KEYWORDS MGC.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS  
1 (bases 1 to 2358)  
Strausberg,R.  
TITLE Direct Submission  
JOURNAL Submitted (23-JAN-2003) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Iohar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMN)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnarathne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H.,  
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,  
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>  
Series: IRAK Plate: 6 Row: K Column: 9

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency OKF analysis, Genomescan gene prediction, Similarity but not identity to protein.

FEATURES  
source  
Location/Qualifiers

1..2358  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="129, C57BL/6J, FVB/N"  
/db\_xref="taxon:10090"  
/clone="MGC:77003 IMAGE:3155470"  
/tissue\_type="mammary tumor. Brcal-/fl; MMTV-Cre model. 10 months old, gross tissue."  
/clone\_1ib="NCI CGAP Mam3"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
136..1023  
/codon\_start=1  
/product="cyclin D1"  
/protein\_id="AAH44641.1"  
/db\_xref="GI:27924089"  
/translation="MEHQLCCCEVETIRRAYPTNLINDRVLRAMKTEKTCAPSVY  
/translation="MEHQLCCCEVETIRRAYPTNLINDRVLRAMKTEKTCAPSVY  
FKVQKEIVPSMKRIIVATWMLVECEKCEEEVPLAMNYLDRLSLERKSRLOL  
GATCMFVPSMKRIIVATWMLVECEKCEEEVPLAMNYLDRLSLERKSRLOL  
IEHFLSKMPEADENKQIRKIAQTFVACATDVAFISNPSMVAAGSVAAAGNLG  
SPNNFLSCYRTTHFLSRVTKCDPDLRACQEQLEALLSSLRQAQNVDPKATEEGE  
VEEAGLACTPTDVRDVI"

BASE COUNT 611 a 587 c 617 g 543 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 2358;  
Best Local Similarity 95.0%; Pred. No. 2.1e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AATGTTCTTAAGATTTT 20  
1438 AATGTTCTTAAGATTTT 1419

RESULT 5  
AX695351/c 3737 bp DNA linear PAT 31-MAR-2003  
LOCUS Sequence 978 from Patent WO03008583.  
DEFINITION AX695351  
ACCESSION AX695351.1 GI:29418501  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Sagres Discovery (US)  
Morris, D.W. and Engelhard, E.K.  
Novel compositions and methods for cancer  
Patent: WO 03008583-A 978 30-JAN-2003;

FEATURES  
source  
Location/Qualifiers  
1..3737  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"

BASE COUNT 921 a 904 c 988 g 924 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 3737;  
Best Local Similarity 95.0%; Pred. No. 1.9e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20  
1440 AATGTTCTTAAGATTTT 1421

RESULT 6

S78355 3737 bp mRNA linear ROD 21-SEP-1995  
LOCUS Cyl-1-cyclin D1 [mice, BALB/c, brain, mRNA, 3737 nt].  
DEFINITION S78355  
ACCESSION S78355  
VERSION S78355.1 GI:994896

KEYWORDS

SOURCE

Mus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
GenBank staff at the National Library of Medicine created this  
entry [NCBI glibseq 167715] from the original journal article.  
This sequence comes from Fig. 1.  
Map location: 7.  
Location/Qualifiers

FEATURES  
source  
Location/Qualifiers  
1..3737  
/organism="Mus sp."  
/mol\_type="mRNA"  
/db\_xref="taxon:10095"

gene  
1..3737  
/gene="Cyl-1"  
/note="cyclin D1"  
138..1025  
/gene="Cyl-1"  
/note="This sequence comes from Fig. 1"

CDS

BASE COUNT 921 a 904 c 988 g 924 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 3737;  
Best Local Similarity 95.0%; Pred. No. 1.9e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20  
1440 AATGTTCTTAAGATTTT 1421

RESULT 7

RATCYCLD1/c 3751 bp mRNA linear ROD 09-JAN-2003

LOCUS RATCYCLD1  
DEFINITION Rattus rattus mRNA for cyclin D1, complete cds.  
ACCESSION D14014  
VERSION D14014.1 GI:413908

KEYWORDS

SOURCE

ORGANISM

Rattus rattus (black rat)  
Rattus rattus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS  
TITLE  
1 (bases 1 to 3751)  
Tamura, K., Kanaoka, Y., Jimo, S., Nagata, A., Ogiso, Y., Shimizu, K.,  
Hayakawa, T., Nojima, H. and Okayama, H.  
Cyclin G: a new mammalian cyclin with homology to fission yeast  
Cyclin G.

JOURNAL  
MEDLINE  
Oncogene 8 (8), 2113-2118 (1993)  
9330551

PUBMED 8336937  
 REFERENCE 2 (bases 1 to 3751)  
 AUTHORS Ogiso, Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-JUN-1993) Yaku Ogiso, Research Development Corporation of Japan, Okayama Cell Switching Project, Pasteur Building 4F, 103-5 Tanaka Monzen-cho, Sakyo-ku, Kyoto, Kyoto 606, Japan (Tel:81-75-712-5406, Fax:81-75-712-5492)

FEATURES  
 Source Location/Qualifiers  
 1..3751

/organism="Rattus rattus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10117"  
 /cell\_line="NRK-49P"  
 /cell\_type="fibroblast"  
 /tissue\_type="kidney"  
 /clone\_lib="pcd2 NRK CDNA library"  
 126..1013  
 /codon\_start=1  
 /product="cyclin D1"  
 /protein\_id="BAA03115.1"  
 /db\_xref="GI:413909"  
 /translation="MEHQLCCEVETIRAYDPDTLNRPGRLAMLTETCAVS  
 FKQVREIVPSMRKIVATWMLVEGECSEEVPLANNYIDRLSEPLKKSRLQ  
 GATCMFVASKKETITLTAETLCYDTSIRPEELQWELLVKKMLAAWTPHF  
 IEHFSKVPADENKQIRKQAQTFVALCAVDKFIENPSPMAAGSVVAMQNLG  
 SPNNLSCTRTTTSRYTKCDPDLACQEIHALLESSLRQMQNIDPKATBECE  
 VEEBAGLACTTTRDVIDI"

BASE COUNT 873 a 914 c 1004 g 960 t  
 ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 3751;  
 Best Local Similarity 95.0%; Pred. No. 1.9e+03;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CY 1 AATGTTCTTAAGATTTT 20  
 Db 1427 AATGTTCTTAAGATTTT 1408

RESULT 8 5019 bp DNA linear INV 04-FEB-1999  
 LOCUS AF060234/c  
 DEFINITION Vairimorpha necatrix largest subunit of RNA polymerase II (RPB1)  
 ACCESSION AF060234  
 VERSION AF060234.1 GI:4001823  
 KEYWORDS  
 SOURCE Vairimorpha necatrix  
 ORGANISM Vairimorpha necatrix  
 Eukaryota; Fungi; Microsporidia; Burenellidae; Vairimorpha.  
 REFERENCE 1 (bases 1 to 5019)  
 Hirt, R.P., Logsdon, J.M. Jr., Healy, B., Dorey, M.W., Doolittle, W.F.  
 Microsporidia are related to Fungi: evidence from the largest  
 subunit of RNA polymerase II and other proteins  
 Proc. Natl. Acad. Sci. U.S.A. 96 (2), 580-585 (1999)  
 9692676  
 2 (bases 1 to 5019)  
 Hirt, R.P. and Healy, B.  
 Direct Submission  
 JOURNAL Submitted (16-APR-1998) Zoology, The Natural History Museum,  
 Cromwell Road, London SW7 5BD, UK

FEATURES  
 Source Location/Qualifiers  
 1..5019  
 /organism="Vairimorpha necatrix"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:6039"  
 101..4918  
 /gene="RPB1"  
 101..4918  
 /gene="RPB1"

/note="RP021"  
 /codon\_start=1  
 /product="largest subunit of RNA polymerase II"  
 /protein\_id="AADI2604.1"  
 /db\_xref="GI:4001824"  
 /translation="MFDEIVKRISSIOFGIPEPEIRKSSVQIHPETWENGPPKS  
 GCLIDLKMGTERAFILCSSCEKNPSPGPHGHGILETKPMHVGWTKIKKILGCVF  
 YCSRLKISTKRLKDLNPNVNISTKVSAGGEIGENPTCGNQAPVTKKSGMLAF  
 PKGEESGKVIINGERVNLIKTVNEDVIFGDFQDFKPEMLILTVLVPPSVR  
 PSIMWEGMLPABDITLKLADIVYANLKYELGAGVADVDEQLQPHIATMD  
 NDISGQPAIOLKSGRPLKISARLKGKGVNKLQKRVDFASBPVYAD  
 IGVPSRLAKIHTPEIITPPNIDRLTLVNGENGVNAGVTRFANQGRIDLPNYSBE  
 IKBEGVVERHMOQDGVVLPNROPGLHKMSMAHVRVNEGTFPLNLSVSPYAD  
 FDDENMLHMPQSYNSKALEELCLVSKVLSPOSKPVGIVQDSTLRLFTTDS  
 FPRRETMQLLYGVNINNVETDSKTLTHDSFNNLHTESSNMKIMLNEPAISY  
 PKLMTGKQILSYILPNTLYNGKSNHNEDELVNDSYVIRNGEILSGIDKAVG  
 STGGGLIHIANDEGDRVTCFDDQKRMNLYFATINAFSIGIDALDKETMGQV  
 RSJETAKBOVNETIYVAKQKRLRLPGMSRESFESQVYILNKPRDISGASASLS  
 FCNNRTWTLASGKSFINI SOLVACLQGNVEGRIPGEGVRLPHSKRDYSGKS  
 RGVENSIVKGIPEEFPHANGREGIDIAIKTAEGYIORPLVYKAMEDATVLD  
 SVRGADGTYOYEGEDGFDAFLMEQKTHDVAKTQDVSFKNLHDMFTDNLDR  
 KENAVDQYKLLTVDVNLQKILYDFEMLNENYKYEKQVSNLFCNFORINLAIVK  
 FCORGDLSPLVILDTLKNLEMLPIKULIEILIKYLSIKRLINELVSTENWI  
 LKEIKFKILSIISPNEMVGTLANQSVBPAQMLNLHAGVSNANTMTPVRLKEI  
 INVAKNIKTPCKKIKYKDPKNTLQSELEFSDIKSLCEFSSEIYYOPIVEDTS  
 IKEDDPOEYVDFPDEHDFSKMFKTIRIKRILKSLKMLNKLKLVNKLHAFRN  
 IPIIIRSDENQNLIRIKRILSIISNNVENVNLQVKNILKMLKLVNKLHAFRN  
 DKEWYLOTQDVCIREIFSHPNVGHVLTNDNLNIEVIGIEARETIINELTVID  
 QNGSVNHRHLSLADVMTKGYLGLTIRHGVNKGCTGRASFETIIVLNLVLD  
 ARKYTKGTENIMMGLAPLGTGIGNLIDVSLDLKAIPLSKREYVEVDLDAV  
 PVSENLSSISGNMSPAYLVEGNRYAPRTSLSPSPSPSPSPSPSPSPSPSP  
 YSP  
 PTPSP

BASE COUNT 1874 a 721 c 909 g 1515 t  
 ORIGIN

Query Match 92.0%; Score 18.4; DB 3; Length 5019;  
 Best Local Similarity 95.0%; Pred. No. 1.7e+03;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CY 1 AATGTTCTTAAGATTTT 20  
 Db 452 AATGTTCTTAAGATTTT 433

RESULT 9 5226 bp DNA linear HTG 16-NOV-1999  
 LOCUS AC014915  
 DEFINITION Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
 ACCESSION AC014915  
 VERSION AC014915.1 GI:6436420  
 HTG: HTGS PHASE2.  
 KEYWORDS  
 SOURCE Drosophila melanogaster (fruit fly)  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE 1 (bases 1 to 5226)  
 Adams, M. and Venter, J.C.  
 Direct Submission  
 JOURNAL Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD, USA

COMMENT  
 This sequence was identified as CDW:10212590 by the submitter.  
 For further information on this sequence e-mail to fly@celera.com.  
 \* NOTE: This is a 'working draft' sequence.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 Location/Qualifiers  
 1..5226  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"

BASE COUNT	1627 a	964 c	1010 g	1625 t	
ORIGIN					Length 5226;
Query Match					92.0%; Score 18.4; DB 2;
Best Local Similarity					95.0%; Pred. No. 1.7e+03;
Matches	19; Conservative	0; Mismatches	1; Indels	0; Gaps	0;
QY	1 AATGTTCTTAGATTGTTT 20				
Db	2097 AATGTTCTTAGATTGTTT 2116				
RESULT 10					INV 29-AUG-1995
DMHISH333	5854 bp DNA linear				
LOCUS	Drosophila melanogaster genes for histone H3.3 and OSt48.				
DEFINITION	X81207.1 GI:9633027				
ACCESSION	X81207				
VERSION	histone H3.3: oligosaccharyltransferase 48 kDa subunit; oSt48 gene				
KEYWORDS	Drosophila melanogaster (fruit fly)				
SOURCE	Drosophila melanogaster				
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 Akhmanova, A.S., Bindels, P.C., Xu, J., Miedema, K., Kremer, H. and Henning, W.				
AUTHORS	Structure and expression of histone H3.3 genes in Drosophila melanogaster and Drosophila hydei				
JOURNAL	Genome 38 (3), 566-600 (1995)				
MEDLINE	96023949				
PUBMED	7557364				
REFERENCE	2 (bases 1 to 5854)				
AUTHORS	Akhmanova, A.S.				
TITLE	Direct Substitution				
JOURNAL	Submitted (03-SEP-1994) A.S. Akhmanova, Catholic University of Nijmegen, Dept of Molecular & Developmental, Genetics, Toernooiveld, 6525 BD Nijmegen, NETHERLANDS				
FEATURES	Location/Qualifiers				
source	1..5854				
	/organism="Drosophila melanogaster"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:7227"				
	/map="9D"				
	/clone="DHE7.2"				
	/clone_1b="DASH11"				
	1..993				
	/gene="OSt48"				
	<1..993				
	/gene="OSt48"				
	/codon_start=1				
	/product="Oligosaccharyltransferase 48kDa subunit"				
	/protein_id="CAA57079.1"				
	/db_xref="GI:963028"				
	/db_xref="FLYBASE:FBgn0014868"				
	/db_xref="SWISS-PROT:Q24319"				
	/translation="EPASCEGFELDENNAVIDLHLDHYVSDAGDDTTLISAKKLRTIVKANRQRDRPLIRGCGTILADKENPLVLLTAESTAYSNPEASVDYFRLRVIALQANNAARVVFSGSLFESDESGVAFHHLAQRNVASSWFGETGQGVASVQHNKEGELHFTDQAYITTPPVYITGIELVQGEARAFKASLEVRIDPEVRYVLLKQNTGAYQAKFKIPVGVGVQPKVDYRVGYTHLYSTVOVPLEHGYERFIPSAFPYTSAFSMWIGVFVFLHFDPEVGAAGKDKKSO				
gene	1335..5854				
	/gene="histone H3.3"				
MRNA	Join(1335..1420,3304..3609,3749..>5854)				
	/gene="histone H3.3"				
exon	1335..1420				
	/gene="histone H3.3"				
	/number=1				
intron	1421..3303				
	/gene="histone H3.3"				
	/number=1				
exon	3304..3609				

CDS	/gene="histone H3.3"
	/number=2
	join(3)28..3609,3749..3877)
	/gene="histone H3.3"
	/codon_start=1
	/protein_id="CAAS7080.1"
	/db_xref="GI:963029"
	/db_xref="FLYBASE:FBgn0004828"
	/db_xref="SWISS-PROT:P06351"
	/translation="MARTKOTARKSTGKAIPRKQATLKAAKSAPSGGVKKPHRRYP GTVALPERRRYSOKTELLIRLPFORLVRRIADFTIDLFPSAIALGAQEASEAVLV GLFEDTNICATHARVITMPKDIOAIRINGERBA"
Intron	3610..3748
	/gene="histone H3.3"
	/number=2
exon	3749..3854
	/gene="histone H3.3"
	/number=3
	4099
polyA_site	/gene="histone H3.3"
	4104
polyA_site	/gene="histone H3.3"
	4117
polyA_site	/gene="histone H3.3"
	4131
polyA_site	/gene="histone H3.3"
	4144
polyA_site	/gene="histone H3.3"
	4742
polyA_site	/gene="histone H3.3"
	4753
	/gene="histone H3.3"
BEGIN COUNT	1644 a 1314 c 1280 g 1616 t
ORIGIN	
Query Match	92.0%; Score 18.4; DB 3; Length 5854;
Best Local Similarity	95.0%; Pred. No. 1.7e+03; Indels 0; Gaps 0;
MATCHES	19; Conservative 0; Mismatches 1;
OY	1 AATGTTCTTAAGATTTC T 20
Dn	2691 AATGTTCTTAAGATTTC T 2710
RESULT 11	
AJ344941	6164 bp DNA linear PAT 01-FEB-2002
LOCUS	AJ344941 Sequence 12 from Patent WO0200928.
DEFINITION	AJ344941
ACCESSION	AJ344941 GI:18492827
VERSION	AJ344941.1
KEYWORDS	'synthetic construct'
SOURCE	'artificial sequences.'
ORGANISM	
REFERENCE	1 Oler,A., Piepenbrock,C. and Berlin,K. Diagnosis of diseases associated with the immune system Patent: WO 0200928-A 12 03-JAN-2002; EpiGenomics AG (DE) Location/Qualifiers 1..6164 /organism="Synthetic Construct" /mol_type="genomic DNA" /db_xref="taxon:32630" /notes="Chemically treated genomic DNA (Homo sapiens)" 1 others
TITLE	
JOURNAL	
FEATURES	
source	
BASE COUNT	1716 a 134 c 1458 g 2855 t
ORIGIN	
Query Match	92.0%; Score 18.4; DB 6; Length 6164;
Best Local Similarity	95.0%; Pred. No. 1.7e+03; Indels 0; Gaps 0;
MATCHES	19; Conservative 0; Mismatches 1;
OY	1 AATGTTCTTAAGATTTC T 20

Db 4643 AATGTTTCTTAAGATTTTT 4662

RESULT 12  
AC014426/C  
LOCUS AC014426 16243 bp DNA linear HTG 16-NOV-1999  
DEFINITION Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*.  
AC014426  
AC014426.1 GI:6436909  
HTG: HTGS\_PHASE2.  
KEYWORDS Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 16243)  
AUTHORS Adams, M. and Venter, J.C.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,  
Rockville, MD, USA  
COMMENT This sequence was identified as CDM:10210291 by the submitter.  
For further information on this sequence e-mail to fly@celera.com.  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
LOCATION/Qualifiers  
1. 16243  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"

BASE COUNT 5707 a 3127 c 2757 g 4652 t

Query Match 92.0%; Score 18.4; DB 2; Length 16243;  
Best Local Similarity 95.0%; Fred. No. 1.3e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 AATGTTTCTTAAGATTTTT 20  
3760 AATGTTTCTTAAGATTTTT 3741

RESULT 13  
AX695350 28486 bp DNA linear PAT 31-MAR-2003  
LOCUS AX695350  
DEFINITION Sequence 977 from Patent WO03008583.  
AC014426  
AX695350.1 GI:29418500  
KEYWORDS Mus musculus (house mouse)  
Mus musculus  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
ORGANISM

REFERENCE  
AUTHORS Morris, D.W. and Engelhardt, E.K.  
TITLE Novel compositions and methods for cancer  
JOURNAL Patent: WO 03008583-A 977 30-JAN-2003;  
Sagres Discovery (US)  
FEATURES  
source 1. 28486  
Location/Qualifiers

BASE COUNT 6341 a 6700 c 6768 g 6414 t 2263 others  
ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 28486;  
Best Local Similarity 95.0%; Fred. No. 1.1e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 AATGTTTCTTAAGATTTTT 20

Db 15914 AATGTTTCTTAAGATTTTT 15895

RESULT 14  
CEK10H10 30033 bp DNA linear INV 21-MAY-2003  
DEFINITION Caenorhabditis elegans cosmid K10H10, complete sequence.  
Z83236  
Z83236.1 GI:1729646  
HTG.  
KEYWORDS Caenorhabditis elegans  
Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
Rhabditidae; Rhabditidae; Pelodetidae; Caenorhabditis.  
1  
none.  
REFERENCE  
AUTHORS none.  
TITLE  
JOURNAL Genome sequence of the nematode C. elegans: a platform for  
MEDLINE investigating biology. The C. elegans Sequencing Consortium  
PUBMED Science 282 (5396), 2012-2018 (1998)  
99069613  
9851916  
REMARK The C. elegans Sequencing Consortium.  
2 (bases 1 to 30033)  
Percy, C.M.  
AUTHORS Direct Submission  
JOURNAL Submitted (09-DEC-1996) Nematode Sequencing Project, Sanger  
Institute, Hinxton, Cambridgeshire CB10 1SA, England and Department of  
Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:  
jes@wustl.edu  
Coding sequences below are predicted from computer analysis, using  
predictions from Genefinder (P. Green, U. Washington), and other  
available information.  
Current sequence finishing criteria for the C. elegans genome  
unambiguously on both strands, or on a single strand with both a  
dye primer and dye terminator reaction, from distinct subclones.  
Exceptions are indicated by an explicit note.  
IMPORTANT: This sequence is not the entire insert of clone K10H10.  
It may be shorter because we only sequence overlapping sections  
once, or longer because we arrange for a small overlap between  
neighbouring submissions.  
The true left end of clone K10H10 is at 7962 in  
sequence Z83110.  
The true right end of clone F57C2 is at 29930 in this sequence. The  
start of this sequence (1. 104) overlaps with the end of sequence  
Z81468.  
The end of this sequence (29930. 30033) overlaps with the start of  
sequence Z83110.  
For a graphical representation of this sequence and its analysis  
see: - [http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?](http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=K10H10)  
name=K10H10  
IMPORTANT: This sequence is NOT necessarily the entire insert of  
the specified clone. It may be shorter because we only sequence  
overlapping sections once, or longer because we arrange for a small  
overlap between neighbouring submissions.  
LOCATION/Qualifiers  
1. 30033  
/organism="Caenorhabditis elegans"  
/mol\_type="genomic DNA"  
/strain="Bristol N2"  
/db\_xref="taxon:6239"  
/chromosome="II"  
/clone="K10H10"  
join(complement(1..212), complement(Z81468.1:23725..24037),  
complement(Z81468.1:22690..23349),  
complement(Z81468.1:21450..21734),  
/gene="usp-14"  
join(complement(1..212), complement(Z81468.1:23725..24037),  
complement(Z81468.1:22690..23349),  
complement(Z81468.1:21450..21734))

FEATURES  
source

gene  
CDS



```

/gene="usp-14"
/standard_name="C13B4.2"
/notes="contains similarity to Pfam domain: PF00442
(ubiquitin carboxyl-terminal hydrolases family 2),
Score=31.3, E-value=3.1e-06, N=1, PF00443 (ubiquitin
carboxyl-terminal hydrolase family 2), Score=108.6,
E-value=4e-29, N=1
CDNA EST cm1f1 comes from this gene; CDNA EST Yk110f1.5
comes from this gene
CDNA EST Yk119a8.3 comes from this gene; CDNA EST
Yk119a8.5 comes from this gene
CDNA EST Yk143a12.3 comes from this gene; CDNA EST
Yk173f1.3 comes from this gene
CDNA EST Yk187b6.3 comes from this gene; CDNA EST
Yk143a12.5 comes from this gene
CDNA EST Yk173f1.5 comes from this gene; CDNA EST
Yk187b6.5 comes from this gene
CDNA EST Yk190b4.5 comes from this gene; CDNA EST
Yk190b4.3 comes from this gene
CDNA EST Yk455e2.5 comes from this gene; CDNA EST
Yk191g6.3 comes from this gene
CDNA EST Yk269a1.3 comes from this gene; CDNA EST
Yk282f4.3 comes from this gene
CDNA EST Yk283f6.3 comes from this gene; CDNA EST
Yk364h11.3 comes from this gene
CDNA EST Yk391b5.3 comes from this gene; CDNA EST
Yk191g6.5 comes from this gene
CDNA EST Yk269a1.5 comes from this gene; CDNA EST
Yk282f4.5 comes from this gene
CDNA EST Yk283f6.5 comes from this gene; CDNA EST
Yk364h11.5 comes from this gene
CDNA EST Yk391b5.5 comes from this gene; CDNA EST
Yk619d1.3 comes from this gene
CDNA EST Yk623a10.3 comes from this gene; CDNA EST
Yk619d1.5 comes from this gene
CDNA EST Yk623a10.5 comes from this gene; CDNA EST
Yk770e04.3 comes from this gene
CDNA EST Yk832e05.3 comes from this gene; CDNA EST
Yk832e05.5 comes from this gene
CDNA EST Yk1061a11.5 comes from this gene; CDNA EST
Yk1061a11.3 comes from this gene
CDNA EST Yk118b05.5 comes from this gene; CDNA EST
Yk118b05.3 comes from this gene
CDNA EST U3223 comes from this gene; CDNA EST Yk1278e05.3
comes from this gene
CDNA EST Yk1290c12.5 comes from this gene; CDNA EST
Yk1290c12.3 comes from this gene"
/codon_start=1
/product="C. elegans USP-14 protein (corresponding
sequence C13B4.2)"
/protein_id="CAB05785.1"
/db_xref="GI:3878545"
/db_xref="GOA:Q17361"
/db_xref="SWISS-PROT:Q17361"
/translation="MPIVAVKQEKRYVEVDTPASPMVFKALFALQVVEREKVV
IMGRIGDDMEGRTIKENMTIMMGVSEIPEPTVLEKQANDKQAEISALYPC
GLANGTCVFNESCVMQKKEVIELVKAPBEKRIREHNDRLCNATLAFNSRJDORA
LRSKKEPKRPAAILTLSDSPROKFKQDNECLVAGMSVTRITYGSGNITSLR
RIQTETMKCESDEVSEKRYERNQILCYVQDPRFLOTGKAFPEEMRNSRLN
RDAAQKNTQISRLPKYLTINIRPFYKSTYTNKALIKLSVQFPMQLDYDICSQELK
DKGLITHKGRSSODGHVAVAMRSSDDGWRFLFDESHVTVVEEA1LKTKSGGDWHS4
VYLVEARYIKQFPELPAPVPTVEAADAFAEVESEKQ"
/locus(1338..1390,1450..1576,1641..1727,1778..1997,
2359..2561,3265..3739,3923..4095)
/gene="K10H10.1"
/standard_name="K10H10.1"
/notes="contains similarity to Pfam domain: PF00083 (Sugar
( and other) transporter), Score=89.7, E-value=0.00037,
N=1

```

## CDS

## gene

```

CDNA EST Yk503h12.5 comes from this gene
CDNA EST Yk653a11.3 comes from this gene
CDNA EST Yk653a11.5 comes from this gene"
/codon_start=1
/product="Hypothetical protein K10H10.1"
/protein_id="CAB05777.1"
/db_xref="GI:3878537"
/db_xref="GOA:Q45678"
/db_xref="SPTREMBL:Q45678"
/translation="MLKRGVEYPPGLKIMTRASHRMWITTFSGTCVLYASRLPI
SAAVAVKERRAMKTDGCVLILSCFPGVATLQVPAIRADYGAELIPYSLSMTMIT
FETPLPDRAYNTNYPLVVLIRITLQCOAHPISLAVSKHLAADKGRVFGIV
LAGSHWGTVALAGISILIEWIGKRLRFQFVGISILWCVPFWVNDKAGRGSRSP
LPDEEVLPHEKDDITIESSHAAATSPCEPWPGLFRHPAFVAAVAQDTGNSYSILFN
MPSYHEPTEPFAKGVVNVPSLAVVLSIVAPVASRLSSEKVTYTRKIMEGAS
LIGAIQCLMVPMTSSFMISLILFTMAARGLHGGVSNPHDFAPNHAQSVFVFN
ACGAILFGVGVYIAGHILEKTNMNSVFVTAQCVCVGMVYTLGTGQK11"
/complement(join(4481..4583,4938..5172,5274..5949))
/gene="K10H10.2"
/complement(join(4481..4583,4938..5172,5274..5949))
/standard_name="K10H10.2"
/notes="contains similarity to Pfam domain: PF00291
(Pyridoxal-phosphate dependent enzyme), Score=428.1,
E-value=2.6e-125, N=1
CDNA EST Yk115e1.5 comes from this gene; CDNA EST
Yk115e1.3 comes from this gene
CDNA EST Yk156a12.3 comes from this gene; CDNA EST
Yk173c5.3 comes from this gene
CDNA EST Yk186f5.3 comes from this gene; CDNA EST
Yk156a12.5 comes from this gene
CDNA EST Yk173c5.5 comes from this gene; CDNA EST
Yk186f5.5 comes from this gene
CDNA EST Yk290d2.3 comes from this gene; CDNA EST
Yk334g3.3 comes from this gene
CDNA EST Yk393g6.3 comes from this gene; CDNA EST
Yk418f11.3 comes from this gene
CDNA EST Yk425c10.3 comes from this gene; CDNA EST
Yk432b7.3 comes from this gene
CDNA EST Yk450f7.3 comes from this gene; CDNA EST
Yk461h2.3 comes from this gene
CDNA EST Yk466c5.3 comes from this gene; CDNA EST
Yk481e1.3 comes from this gene
CDNA EST Yk481e10.3 comes from this gene; CDNA EST
Yk410h8.3 comes from this gene
CDNA EST Yk483d3.3 comes from this gene; CDNA EST
Yk290d2.5 comes from this gene
CDNA EST Yk334g3.5 comes from this gene; CDNA EST
Yk393g6.5 comes from this gene
CDNA EST Yk418f11.5 comes from this gene; CDNA EST
Yk425c10.5 comes from this gene
CDNA EST Yk432b7.5 comes from this gene; CDNA EST
Yk461h2.5 comes from this gene
CDNA EST Yk450f7.5 comes from this gene; CDNA EST
Yk481e1.5 comes from this gene
CDNA EST Yk481e10.5 comes from this gene; CDNA EST
Yk466c5.5 comes from this gene
CDNA EST Yk410h8.5 comes from this gene; CDNA EST
Yk483d3.5 comes from this gene
CDNA EST Yk519h5.3 comes from this gene; CDNA EST

```

```

Query Match          92.0%; Score 18.4; DB 3; Length 3003;
Best Local Similarity 95.0%; Pred. No. 1.1e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 AATGTTCTTAGATTGTTT 20
Db      14322 AATGTTCTTAATATTTT 14341

```

```

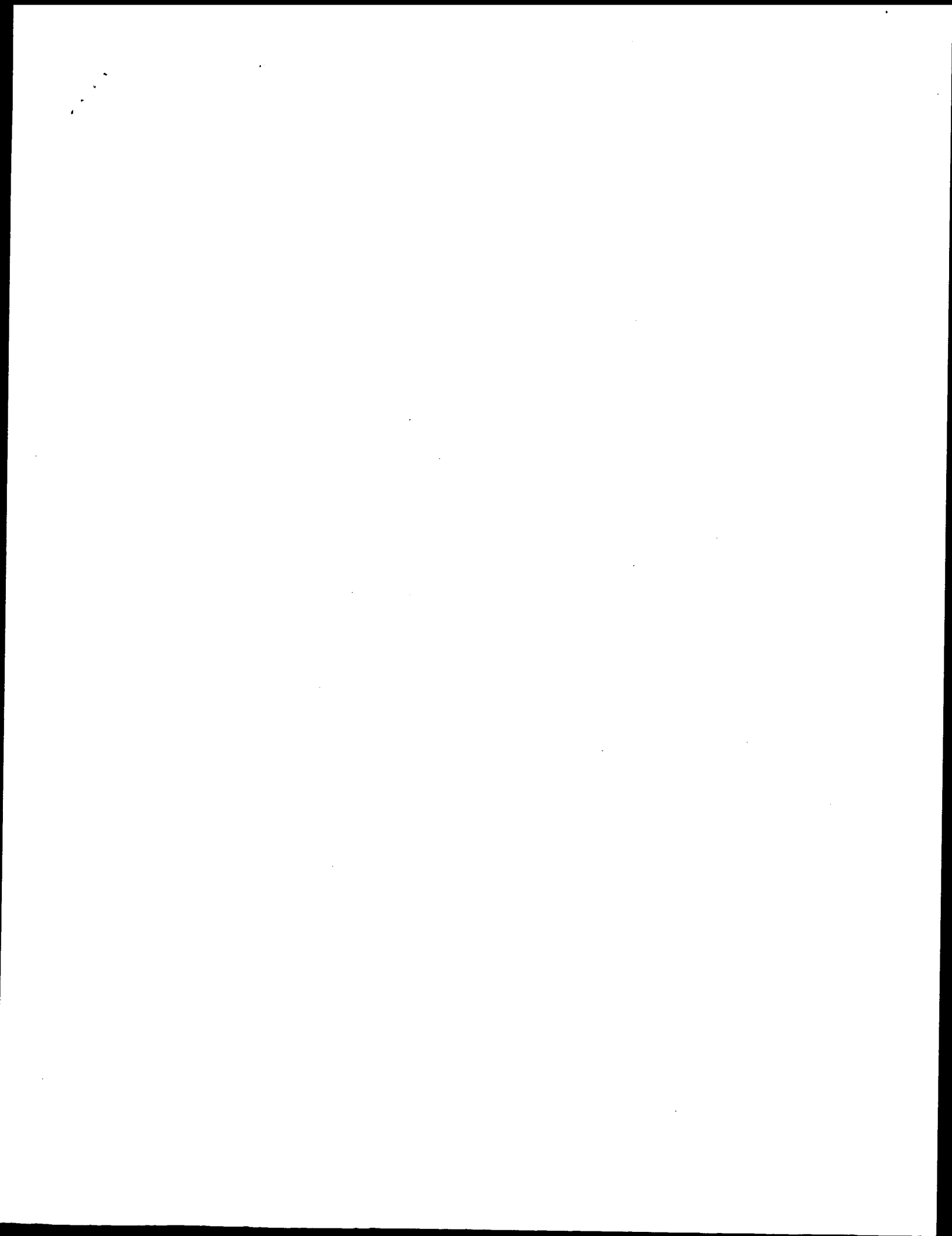
RESULT 15
CEK04H4/c      33930 bp      DNA      linear      INV 21-MAY-2003
LOCUS          CEK04H4

```

DEFINITION Caenorhabditis elegans cosmid K04H4, complete sequence.  
 ACCESSION Z27078  
 VERSION Z27078.3 GI:25004984  
 KEYWORDS HTG.  
 SOURCE Caenorhabditis elegans  
 ORGANISM Caenorhabditis elegans  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
 Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 REFERENCE 1  
 none.  
 AUTHORS Genome sequence of the nematode *C. elegans*: a platform for  
 TITLE investigating biology. The *C. elegans* Sequencing Consortium  
 JOURNAL Science 282 (5396), 2012-2018 (1998)  
 MEDLINE 99069613  
 PUBMED 9851916  
 REMARK The *C. elegans* Sequencing Consortium.  
 REFERENCE 2 (bases 1 to 33930)  
 AUTHORS Ainscough, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (04-NOV-1993) Nematode Sequencing Project, Sanger  
 Institute, Hinxton, Cambridge CB10 1SA, England and Department of  
 Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:  
 jes@sanger.ac.uk or r@nematode.wustl.edu  
 On Nov 14, 2002 this sequence version replaced gi:19571633.  
 Coding sequences below are predicted from computer analysis, using  
 predictions from GeneFinder (P. Green, U. Washington), and other  
 available information.  
 Current sequence finishing criteria for the *C. elegans* genome  
 sequencing consortium are that all bases are either sequenced  
 unambiguously on both strands, or on a single strand with both a  
 dye primer and dye terminator reaction, from distinct subclones.  
 Exceptions are indicated by an explicit note.  
 IMPORTANT: This sequence is not the entire insert of clone K04H4.  
 It may be shorter because we only sequence overlapping sections  
 once, or longer because we arrange for a small overlap between  
 neighbouring submissions.  
 The start of this sequence (1..97) overlaps with the end of  
 sequence Z19154.  
 The end of this sequence (33837..33930) overlaps with the start of  
 sequence Z19153.  
 For a graphical representation of this sequence and its analysis  
 see: [http://wormbase.sanger.ac.uk/perl/ce/elegans/seq/sequence?](http://wormbase.sanger.ac.uk/perl/ce/elegans/seq/sequence?name=K04H4)  
 name=K04H4  
 IMPORTANT: This sequence is NOT necessarily the entire insert of  
 the specified clone. It may be shorter because we only sequence  
 overlapping sections once, or longer because we arrange for a small  
 overlap between neighbouring submissions.  
 [020311 d1] T->G at position 3143 based on Thierry-Mieg EST  
 alignments  
 [020311 d1] + C at position 4614 based on Thierry-Mieg EST  
 alignments  
 [020311 d1] - T at position 4618 based on Thierry-Mieg EST  
 alignments  
 [021031 ar2] + G at 6588 position based on EST alignment [021031  
 ar2] - C at position 6808 based on EST alignment.  
 Location/Qualifiers  
 1..33930  
 /organism="Caenorhabditis elegans"  
 /mol\_type="genomic DNA"  
 /strain="Bristol N2"  
 /db\_xref="taxon:6239"  
 /chromosome="III"  
 /clone="K04H4"  
 complement(join(1214..1309,1727..2508,2567..3839,  
 3929..4789,4844..5086,5180..5434,5491..6515,6573..6975,  
 7381..7532,7941..8107,8382..8449))  
 /gene="emb-9"  
 complement(join(1214..1309,1727..2508,2567..3839,  
 3929..4789,4844..5086,5180..5434,5491..6515,6573..6975,  
 7381..7532,7941..8107,8382..8449))  
 /gene="emb-9"  
 /standard\_name="K04H4.1"  
 /note="similar to collagens, contains similarity to Pfam

domain: PF01391 (Collagen triple helix repeat (20  
 copies)), Score=68.9, E-value=4e-204, N=21; PF01413  
 (C-terminal tandem repeated domain in type 4 procollagen),  
 Score=500.5, E-value=4e-147, N=2  
 CDNA EST CERS009P comes from this gene; CDNA EST yk108f8.5  
 comes from this gene  
 CDNA EST yk35b4.5 comes from this gene; CDNA EST yk95b4.3  
 comes from this gene  
 CDNA EST yk90c8.3 comes from this gene; CDNA EST yk90c8.5  
 comes from this gene  
 CDNA EST yk108f8.3 comes from this gene; CDNA EST  
 yk284d3.3 comes from this gene  
 CDNA EST yk374e1.3 comes from this gene; CDNA EST  
 yk425h10.3 comes from this gene  
 CDNA EST yk501e3.3 comes from this gene; CDNA EST  
 yk284d3.5 comes from this gene  
 CDNA EST yk374e1.5 comes from this gene; CDNA EST  
 yk425h10.5 comes from this gene  
 CDNA EST yk654a8.3 comes from this gene; CDNA EST  
 yk501e3.5 comes from this gene  
 CDNA EST yk755c11.3 comes from this gene; CDNA EST  
 yk755c11.5 comes from this gene  
 CDNA EST yk776h06.3 comes from this gene; CDNA EST  
 yk776h06.5 comes from this gene  
 CDNA EST yk878c05.3 comes from this gene; CDNA EST  
 yk878c05.5 comes from this gene  
 CDNA EST yk1029f07.3 comes from this gene; CDNA EST  
 yk1053c02.5 comes from this gene  
 CDNA EST yk1053c02.3 comes from this gene; CDNA EST  
 yk1062a05.5 comes from this gene  
 CDNA EST yk1062a05.3 comes from this gene; CDNA EST  
 yk1068c10.3 comes from this gene  
 CDNA EST yk1068f08.5 comes from this gene; CDNA EST  
 yk1068f08.3 comes from this gene  
 CDNA EST yk1113c11.3 comes from this gene; CDNA EST  
 yk1202a01.3 comes from this gene  
 CDNA EST yk1278a06.5 comes from this gene; CDNA EST  
 yk1278a06.3 comes from this gene  
 CDNA EST yk1285a06.5 comes from this gene; CDNA EST  
 yk1285a06.3 comes from this gene  
 CDNA EST yk1285d05.5 comes from this gene; CDNA EST  
 yk1285d05.3 comes from this gene  
 CDNA EST yk1313a04.5 comes from this gene; CDNA EST  
 yk1313a04.3 comes from this gene  
 CDNA EST yk1025a11.3 comes from this gene  
 /codon\_start=1  
 /product="C. elegans EMB-9 protein (corresponding sequence  
 K04H4.1)"  
 /protein\_id="CAA81584.3"  
 /db\_xref="GI:25004987"  
 /db\_xref="GOA:P17139"  
 /db\_xref="SWISS-PROT:P17139"  
 /translation="MSRLSLGLTAAVVLSSFCQDRTHVDAAMAKGCAAPVCPT  
 KGERGNPFGEPGEPHGPAGQDGPBGAGAEEDFQDMGSKARGGRGLPSG  
 GHPIGLDGLGPKLKEGEGIPGNGTGFPPMPLAGPPGSGNGNGRGLSPG  
 EGVAISOGRKVGSGSGSVPLKNSGNGVGLGPGGPPGGLRPGPGVGLKRGV  
 GRTSGYKGEKGLPGPPGPPGSGSPASPTIMETVLQGVGVAGVAGEGSDPVG  
 PPGMLGLDPPRGYGLKQKQDLDDAQRGSLICNKTKLFLCVNVERTEFLETMEK  
 RDPKENDLELQDTQLREGLNQDQVDSKELYDRKHLKEVLRLHMEIKGS  
 TERFEARDQDMWVSPGVGPRGVPAAGAGQGIQGMPTYTERKGSGEGYFAG  
 EPLGLBEPDQGPBGDGLDGYDIQGPGLDQSGRGLPGLPDIGPVGSGKGP  
 GTGVNKGVPFGMTGLPGSPGMFGRIGVGYGPGNNGEREDGCGYPRGNAGDP  
 GFGPMNGYPPGPGPDGHDGCMGAPGAKRSASDLSGSGRLPGLPGYMGAGAG  
 EIVGPMNPGIPLGLKGDHLPLRGRPSGDLGLGYGCGPQNGFPLGLQSPGLAGID  
 KGRKROSLGIPGLQSPGSGSPFLPGLPFGKSGAGDGLPGLGAGAPRQIPALPI  
 VNQVAGQPGVDMGELPFDGRADGLPGLPFGVGDGPGTGERGMDGLPGFPLHAE  
 PKMKQQGEVGFANGIDDCBEPGLDVPAGAPAGAGETGFPQGVGYPGPDHAE  
 AAGLPBGDYPBGDGLPGLTGPYGEAAMNODAPAGQSGSGSLGIDGKGRDGT  
 PGTGRQDGGPGEAGAPGONGDGVPRAGPAGGYGSGDGGYPPSPGIPGDDYV  
 GFLRLREHGDNGLPGLBEGRGSGSLGDGVGYPERHETDGLPGLGAGDQGPYV  
 EAGEPGTGRGQGPGRGNLARYGQPDVYPPGDGPGLPDGLGALNGERDND  
 SYGNPRLSGPBDAGIDGLDGVGPPGPGYGLTGMPLKGSGLPGLPGQNGDIPG  
 QPLBEGCGEDGPPSGPQPGVPGQGRBEGKGYPLIPGNHGLPGLRGQDQGPGLKGE





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:13:22 ; Search time 68.4754 Seconds  
(without alignments)  
7098.748 Million cell updates/sec

Title: US-09-834-291-15  
Perfect score: 20  
Sequence: 1 aatgttcttaagatttttt 20

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estcin.\*  
4: em\_estmu.\*  
5: em\_estrov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hic.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hic.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gse\_hum.\*  
18: em\_gse\_inv.\*  
19: em\_gse\_pln.\*  
20: em\_gse\_vrt.\*  
21: em\_gse\_fun.\*  
22: em\_gse\_man.\*  
23: em\_gse\_mus.\*  
24: em\_gse\_pro.\*  
25: em\_gse\_rtd.\*  
26: em\_gse\_phg.\*  
27: gb\_gse\_vrt.\*  
28: gb\_gse1.\*  
29: gb\_gse2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18.4	92.0	159	9	AV243724 AV243724
C 2	18.4	92.0	160	9	AV174680 AV174680
C 3	18.4	92.0	198	9	AW213875 AW213875
C 4	18.4	92.0	213	9	AV134425 AV134425

C 5	18.4	92.0	226	10	BB834512
C 6	18.4	92.0	237	9	AV133072
C 7	18.4	92.0	240	10	BB048825
C 8	18.4	92.0	255	10	BB546588
C 9	18.4	92.0	270	10	AV135706
C 10	18.4	92.0	278	12	B1289904
C 11	18.4	92.0	280	10	BB049315
C 12	18.4	92.0	283	10	BF334470
C 13	18.4	92.0	298	9	AV219191
C 14	18.4	92.0	299	9	AV032131
C 15	18.4	92.0	328	28	AA027274
C 16	18.4	92.0	332	9	AA229904
C 17	18.4	92.0	339	9	BB546854
C 18	18.4	92.0	355	28	AA055348
C 19	18.4	92.0	360	14	W78658
C 20	18.4	92.0	379	13	BY229322
C 21	18.4	92.0	380	13	BY174614
C 22	18.4	92.0	388	10	BB836659
C 23	18.4	92.0	411	10	BB781889
C 24	18.4	92.0	436	9	AL719273
C 25	18.4	92.0	444	12	BI303131
C 26	18.4	92.0	446	13	BY472560
C 27	18.4	92.0	457	10	BB831513
C 28	18.4	92.0	469	12	BB758495
C 29	18.4	92.0	470	12	BB898746
C 30	18.4	92.0	473	10	AA267268
C 31	18.4	92.0	473	28	BB833441
C 32	18.4	92.0	491	14	CA877820
C 33	18.4	92.0	515	12	B1452813
C 34	18.4	92.0	515	14	CA879995
C 35	18.4	92.0	516	14	CA872437
C 36	18.4	92.0	517	14	CA870300
C 37	18.4	92.0	518	14	CA870688
C 38	18.4	92.0	518	14	CA871824
C 39	18.4	92.0	518	14	CA873313
C 40	18.4	92.0	518	14	CA879757
C 41	18.4	92.0	529	9	AA259960
C 42	18.4	92.0	536	28	AA030889
C 43	18.4	92.0	570	14	CA873049
C 44	18.4	92.0	572	28	BZ175552
C 45	18.4	92.0	572	29	BZ611328

## ALIGNMENTS

RESULT 1  
AV243724/c 159 bp mRNA linear EST 04-NOV-1999  
AV243724 RIKEN full-length enriched, 0 day neonate head Mus  
LOCUS AV243724.1 GI:6231183  
DEFINITION Mus musculus CDNA clone 4831423J17 3' similar to S7835 Cyl-1-cyclin D1  
mRNA sequence.

ACCESSION AV243724  
VERSION AV243724.1 GI:6231183

KEYWORDS EST.  
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

## REFERENCE

1 (bases 1 to 159)  
Kanno, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateo, M., Tomimaga, N., Tsunoda, Y., Watanabe, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.  
Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.  
RIKEN Mouse ESTs (Kanno, H., et al. 1999)  
Contact: Yoshihide Hayashizaki

TITLE  
JOURNAL  
COMMENT



CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNI at:  
 www-bio.llnl.gov/bbrp/image/image.html

MG1:1025833  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 191.

Location/Qualifiers

1.198

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="CZBCH II"

/db\_xref="taxon:10090"

/clone="IMAGE:2645381"

/issue\_type="spontaneous tumor, metastatic to mammary."

Stem cell origin."

/lab\_host="DH10B"

/clone\_lib="NCI-CGAP Lu29"

/note="Organ: lung; Vector: pCMV-Sport6; Site: 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

## BASE COUNT

56 a 38 c 23 g 81 t

## ORIGIN

## Query Match

Best Local Similarity 92.0%; Score 18.4; DB 9; Length 198;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 AATGTTCTTAAGATTCTTT 20

33 AATGTTCTTAAGCTTTT 52

Db

## RESULT 4

## LOCUS

AV134425 213 bp mRNA linear EST 01-JUL-1999  
 AV134425 Mus musculus C57BL/6J 10-11 day embryo Mus musculus cDNA  
 clone 2810006111, mRNA sequence.

## ACCESSION

AV134425

## VERSION

AV134425.1 GI:5320155

## KEYWORDS

EST

## SOURCE

Mus musculus (house mouse)

## ORGANISM

Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 (bases 1 to 213)  
 Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,  
 Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara,  
 A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J.,  
 Kikuchi, N., Kojima, Y., Matsuyama, T., Natsuna, H., Oda, H., Owa, C.,  
 Sato, K., Shibata, Y., Shigenoto, Y., Shiraki, T., Sogabe, Y., Sugahara,  
 Y., Suzuki, H., Suzuki, H., Tateo, M., Tomaru, Y., Tomimaga, N.,  
 Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M.,  
 Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

## AUTHORS

RIKEN Mouse ESTs

## TITLE

Unpublished

## JOURNAL

Genome Science Laboratory

## COMMENT

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

## TITLE

Unpublished

## JOURNAL

Genome Science Laboratory

## COMMENT

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

## TITLE

Unpublished

## JOURNAL

Genome Science Laboratory

## COMMENT

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

## TITLE

Unpublished

## JOURNAL

Genome Science Laboratory

## COMMENT

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

## FEATURES

## source

Location/Qualifiers  
 1.213  
 /organism="Mus musculus"  
 /mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="2810006111"

/sex="mixed"

/dev\_stage="10-11 day embryo"

/clone\_lib="Mus musculus C57BL/6J 10-11 day embryo"

## BASE COUNT

61 a 22 c 54 g 76 t

## ORIGIN

## Query Match

Best Local Similarity 92.0%; Score 18.4; DB 9; Length 213;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 AATGTTCTTAAGATTCTTT 20

208 AATGTTCTTAAGCTTTT 199

Db

## RESULT 5

## LOCUS

BB834512 226 bp mRNA linear EST 19-NOV-2001  
 BB834512 RIKEN full-length enriched, mammary gland RCB-0527  
 JYG-MC(B) CDNA Mus musculus cDNA clone G930031G12 3', mRNA  
 sequence.

## ACCESSION

BB834512

## VERSION

BB834512.1 GI:17012755

## KEYWORDS

EST

## SOURCE

Mus musculus (house mouse)

## ORGANISM

Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 (bases 1 to 226)  
 Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,  
 Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imocani, K., Ishi,  
 Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T.,  
 Nakamura, M., Nishi, K., Nomura, K., Nunasaki, R., Okazaki, Y., Okido, T.,  
 Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K.,  
 Shibata, K., Shigenaga, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa,  
 A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T.,  
 Warahiki, A., Yasunishi, A., Yamamatsu, M. and Hayashizaki, Y.

## AUTHORS

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
 2001)

## TITLE

Unpublished

## JOURNAL

Contact: Yoshihide Hayashizaki

## COMMENT

Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216

## REFERENCE

Email: genome-resgsc.riken.go.jp,  
 url: http://genome-gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

## AUTHORS

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
 2001)

## TITLE

Unpublished

## JOURNAL

Contact: Yoshihide Hayashizaki

## COMMENT

Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216

## REFERENCE

Email: genome-resgsc.riken.go.jp,  
 url: http://genome-gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

## AUTHORS

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
 2001)

## TITLE

Unpublished

## JOURNAL

Contact: Yoshihide Hayashizaki

## COMMENT

Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216

## REFERENCE

Email: genome-resgsc.riken.go.jp,  
 url: http://genome-gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

## AUTHORS

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
 2001)

## TITLE

Unpublished

## JOURNAL

Contact: Yoshihide Hayashizaki

## COMMENT

Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216

## REFERENCE

Email: genome-resgsc.riken.go.jp,  
 url: http://genome-gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

## AUTHORS

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
 2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES  
source  
Location/Qualifiers  
1..226

ORGANISM="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="G930031G12"  
/tissue\_type="mammary gland"  
/cell\_line="RCB-0527 Jyg-MC(B)"  
/clone\_idb="RCB-0527 Jyg-MC(B) CDNA"  
RCB-0527 Jyg-MC(B) CDNA

BASE COUNT  
81 a 31 c 48 g 66 t

Query Match 92.0%; Score 18.4; DB 10; Length 226;  
Best Local Similarity 95.0%; Pred. NO. 1.5e+04;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20  
|||||  
210 AATGTTCTTAAGCTTTT 191

RESULT 6  
AV133072/c 237 bp mRNA linear EST 01-JUN-1999  
LOCUS AV133072 Mus musculus C57BL/6J 11-day embryo Mus musculus CDNA  
ACCESSION clone 2700095E14, mRNA sequence.  
VERSION AV133072  
KEYWORDS AV133072.1 GI:5308802  
SOURCE EST.

ORGANISM  
Mus musculus (house mouse)

REFERENCE  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 237)

AUTHORS  
Carninci, P., Shibata, K., Ozawa, Y., Kono, H., Itoh, M., Aizawa, K.,  
Akahira, S., Akiyama, T., Fukuda, S., Fukunishi, Y., Funayama, T., Hara  
A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J.,  
Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuna, H., Oda, H., Owa, C.,  
Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara  
Y., Suzuki, H., Suzuki, H., Tateo, M., Tomaru, Y., Tomimaga, N.,  
Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M.,  
Miyamatsu, M., Okazaki, Y. and Hayashizaki, Y.

TITLE  
JOURNAL  
COMMENT  
Unpublished  
Contact: Chie Owa  
Genome Science Laboratory  
RIKEN  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-9145  
Fax: 81-298-36-9098  
Email: genome-res@sc.riken.go.jp

Thermostabilization and thermoactivation of thermolabile enzymes by  
trehalose and its application for the synthesis of full length cDNA  
(Proc. Natl. Acad. Sci. U.S.A. 95(12):520-524 (1998))  
Transcriptional sequencing: A method for DNA sequencing using RNA  
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for  
further details.

FEATURES  
source  
Location/Qualifiers  
1..237

ORGANISM="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="2700095E14"  
/sex="mixed"  
/dev\_stage="11-day embryo"  
/clone\_idb="Mus musculus C57BL/6J 11-day embryo"  
83 a 27 c 49 g 78 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 237;  
Best Local Similarity 95.0%; Pred. NO. 1.5e+04;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20  
|||||  
232 AATGTTCTTAAGCTTTT 213

RESULT 7  
BB048825/c 240 bp mRNA linear EST 25-JUN-2000  
LOCUS BB048825 RIKEN full-length enriched, adult male olfactory bulb Mus  
musculus CDNA clone 643056J05 3' similar to S7835 Cyl-1-cyclin D1  
ACCESSION BB048825.1 GI:8455973  
VERSION BB048825  
KEYWORDS BB048825.1 GI:8455973  
SOURCE EST.

ORGANISM  
Mus musculus (house mouse)

REFERENCE  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 240)

AUTHORS  
Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci  
P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Itoh, M.,  
Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,  
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusabe, M.,  
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,  
Ota, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y.,  
Shigemoto, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y.,  
Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya  
T., Tsunoda, Y., Watanabe, S., Watanabe, S., Yamamura, T., Yamataka, I.,  
Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino  
M., Miyamatsu, M. and Hayashizaki, Y.

TITLE  
JOURNAL  
COMMENT  
Unpublished  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.go.jp,  
URL: <http://genome.gsc.riken.go.jp/>

Thermostabilization and thermoactivation of thermolabile enzymes by  
trehalose and its application for the synthesis of full length  
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh, M., Katsunari, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
Tomaru, Y., Carninci, P., Shibata, K., Ozawa, Y., Miyamatsu, M., Okazaki  
Y. and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
19-44 (1999)  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for  
further details.

FEATURES  
source  
Location/Qualifiers  
1..240

ORGANISM="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="643056J05"  
/sex="male"  
/tissue\_type="olfactory brain"  
/dev\_stage="adult"







REFERENCE	1 (bases 1 to 298)
AUTHORS	Kono, H., Aizawa, K., Akahira, S., Ariyama, J., Carninci, P., Endo, T., Furukawa, Y., Harai, A., Hayatsu, N., Hirozane, T., Hori, F.,



REFERENCE  
AUTHORS

1 (bases 1 to 328)  
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
Hood,L.

Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

JOURNAL  
MEDLINE  
PUBMED

99380589  
10449764

## COMMENT

Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center

University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618  
Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence Tagged Connector

Plate: 2255 row: G column: 19

Class: BAC ends  
High quality sequence stop: 328.  
Location/Qualifiers

FEATURES  
source

1..328

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/clone="plate=2255 Col=19 Row=G"

/sex="male"

/clone\_lib="CIT Approved Human Genomic Sperm Library D"

/note="Organ: sperm; Vector: pBelobAC11; BAC clones in  
E-Coli DH10B"

BASE COUNT 102 a 63 c 50 g 113 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 28; Length 328;  
Best Local Similarity 95.0%; Pred. No. 1.4e+04;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTACGTTTTT 20  
|||||

Db 27 AATGTTCTTACGTTTTT 46  
|||||

Search completed: August 1, 2003, 20:43:07  
Job time : 69.4754 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:32:17 ; Search time 6.69868 Seconds  
(without alignments)  
8059.612 Million cell updates/sec

Title: US-09-834-291-15  
Perfect score: 20  
Sequence: 1 aatgttcttaagattttc 20

Scoring table: IDENTITY NUC  
Gapco 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_19Jun03:\*

1:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	18.4	92.0	645	21	AAA75620
C 2	18.4	92.0	645	24	AA520735
C 3	18.4	92.0	2696	23	ABL22726
C 4	18.4	92.0	3273	23	ABL27256
C 5	18.4	92.0	4839	23	ABL14642
C 6	18.4	92.0	6164	24	ABL32039
C 7	17.4	87.0	149	21	AAC29641
C 8	17.4	87.0	536	24	ABN63544

C 9	17.4	87.0	954	24	ABN71515
C 10	17.4	87.0	996	24	ABN66487
C 11	17.4	87.0	1068	24	ABN66079
C 12	17.4	87.0	1528	20	AA339670
C 13	17.4	87.0	1528	24	ABK84789
C 14	17.4	87.0	2483	23	ABL16936
C 15	17.4	87.0	2531	23	ABL11262
C 16	17.4	87.0	4064	24	ABK84306
C 17	17.4	87.0	4664	23	ABL12292
C 18	17.4	87.0	5519	24	ABK39970
C 19	17.4	87.0	215561	24	ABN71527
C 20	16.8	84.0	20	21	AA288724
C 21	16.8	84.0	313	24	ABN20279
C 22	16.8	84.0	366	22	AA135113
C 23	16.8	84.0	366	25	ABK58101
C 24	16.8	84.0	401	22	AA192787
C 25	16.8	84.0	405	22	AA187334
C 26	16.8	84.0	438	22	AAK56912
C 27	16.8	84.0	439	22	AAK69090
C 28	16.8	84.0	450	25	ABX47860
C 29	16.8	84.0	473	21	AAC22374
C 30	16.8	84.0	486	24	ABN63569
C 31	16.8	84.0	515	22	AA546922
C 32	16.8	84.0	515	24	ABK81700
C 33	16.8	84.0	590	24	ABN63363
C 34	16.8	84.0	993	23	AA533595
C 35	16.8	84.0	1462	24	ABK54435
C 36	16.8	84.0	1608	24	AB215263
C 37	16.8	84.0	1843	24	AB215263
C 38	16.8	84.0	2121	22	AAH14223
C 39	16.8	84.0	2742	21	AAH4257
C 40	16.8	84.0	3481	23	AB106966
C 41	16.8	84.0	3562	22	ABK17959
C 42	16.8	84.0	5252	24	ABL31135
C 43	16.8	84.0	5277	23	ABL18542
C 44	16.8	84.0	5816	24	ABK40012
C 45	16.8	84.0	5937	24	ABN80151

## ALIGNMENTS

RESULT 1  
AAA75620/c  
AAA75620 standard; DNA; 645 BP.

AC AAA75620;  
DT 22-JAN-2001 (first entry)

XX Nucleotide sequence of a human zalphall ligand fragment.

DE zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;  
XX tumorigenesis; leukaemia; hematopoiesis; B cell tumour; ss.

OS Homo sapiens.

XX	Key	Location/Qualifiers
FT	exon	436..513
FT		/*tag= a
FT		/number= 4
FT	intron	514..603
FT		/*tag= b
FT		/number= 4
FT		604..645
FT	exon	/*tag= c
FT		/number= 5
FT		/note= "partial exon sequence"

PN W0200053761-A2.

XX 14-SEP-2000.

XX

Streptococcus poly  
Streptococcus poly  
Streptococcus poly  
Renal cancer assoc  
Human cDNA differe  
Drosophila melanog  
Human cDNA differe  
Drosophila melanog  
Human chemically p  
Streptococcus poly  
Human CD95 recepto  
Human ORFX polyuc  
Human musculocele  
CDNA encoding nove  
Human polynucleoti  
Human polynucleoti  
Human immune/haema  
Bovine EST associa  
Human secreted pro  
Human cancer relat  
Human G protein-co  
CDNA encoding nove  
Human cancer relat  
Helicobacter pylor  
Human BMP2 partia  
Human immune syste  
Arabidopsis thalia  
Human cDNA sequenc  
Human secreted pro  
Drosophila melanog  
Human nervous syst  
Human immune syste  
Drosophila melanog  
Human chemically p  
Human chemically m

PF 09-MAR-2000; 2000WO-US06067.  
XX  
PR 09-MAR-1999; 99US-0264908.  
PR 11-MAR-1999; 99US-0265992.  
PR 01-JUL-1999; 99US-0142013.  
XX  
PA (ZYMO) ZYMOGENETICS INC.  
XX  
PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;  
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;  
XX WPI; 2000-565600/52.  
XX  
PT New human cytokine, designated zalphal1 ligand, useful for stimulating  
PT the proliferation and/or development of haematopoietic cells in vitro  
PT and in vivo, and for treating tumorigenesis -  
XX  
PS Example 52; Page 251; 256pp; English.  
XX  
CC The present sequence represents part of the human zalphal1 ligand gene.  
CC zalphal1 ligand is a cytokine. The zalphal1 ligand is useful for  
CC stimulating the proliferation and development of haematopoietic cells  
CC in vitro and in vivo. Zalphal1 ligand polynucleotides can be used as  
CC primers or probes for cloning the zalphal1 gene. The zalphal1 ligand is  
CC useful for treating tumorigenesis. A zalphal1 ligand-saporin fusion  
CC toxin may be used for treating leukemias and lymphomas. Antagonists  
CC against zalphal1 ligand are useful as research reagents for  
CC characterizing ligand-receptor interaction. Antagonists are also useful  
CC for inhibiting expansion, proliferation, activation and differentiation  
CC of cells involved in regulating hematopoiesis. The zalphal1 ligand may  
CC also be used to stimulate an immune response against B cell tumour, a  
CC virus, a parasite or a bacterium. The zalphal1 polypeptides,  
CC polynucleotides, antagonists, agonists and antibodies are also useful  
CC for the detection, diagnosis, prevention, and treatment of diseases  
CC associated with a zalphal1 ligand genetic defect.  
XX  
SQ Sequence 645 BP; 244 A; 107 C; 89 G; 204 T; 1 other;  
XX  
Query Match 92.0%; Score 18.4; DB 21; Length 645;  
Best Local Similarity 95.0%; Pred. No. 2.7e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AATGTTCTTAAGATTGTTT 20  
DB 383 AATGTTCTTAAGATTGTTT 364

RESULT 2  
AAS20735/c  
ID AAS20735 standard; DNA; 645 BP.  
XX  
AC AAS20735;  
XX  
DT 09-APR-2002 (first entry)  
XX  
DE Human zalphal1 ligand gene partial intron 3-partial exon 5 sequence.  
XX  
KW Cytokine; zalphal1 ligand; zalphal1 receptor; NK cell progenitor;  
KW natural killer cell proliferation; T-cell proliferation;  
KW B-cell proliferation; anti-tumour response; immune system;  
KW immunostimulant; cytoskeletal; human; ds.  
XX  
OS Homo sapiens.  
XX  
PN US6307024-B1.  
XX  
PD 23-OCT-2001.  
XX  
PF 09-MAR-2000; 2000US-0522217.  
XX  
PR 09-MAR-1999; 99US-123547P.  
PR 11-MAR-1999; 99US-123904P.  
PR 01-JUL-1999; 99US-142013P.

XX  
PA (ZYMO) ZYMOGENETICS INC.  
XX  
PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;  
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;  
XX WPI; 2002-040208/05.  
XX  
PT New zalphal1 ligand polypeptides and polynucleotides, useful for  
PT stimulating proliferation, activation, differentiation and/or induction  
PT of inhibition of specialized cell function, or for stimulating an  
PT antigenic response -  
XX  
PS Example 52; Column 187-188; 105pp; English.  
XX  
CC The present invention relates to the isolation of a novel cytokine,  
CC zalphal1 ligand and the polynucleotide encoding it. The invention  
CC also gives the sequence for the zalphal1 receptor and the polynucleotide  
CC encoding it. The zalphal1 ligand polypeptide stimulates proliferation of  
CC natural killer (NK) cells or NK cell progenitors, the activation of NK  
CC cells, proliferation of T-cells, proliferation of B-cells stimulated  
CC with anti-CD40 antibodies, stimulates an antigenic response in a mammal,  
CC and reduces proliferation of B-cells stimulated with anti-1gM antibodies.  
CC The zalphal1 ligand polypeptide is also useful in preparing antibodies  
CC that bind to zalphal1 ligand epitopes. The zalphal1 ligand  
CC polynucleotides can be used as probes or primers to clone regions  
CC of a zalphal1 ligand gene, and in gene therapy. Zalphal1 ligand may  
CC also be used to identify inhibitors of its activity, to enhance the  
CC generation of anti-tumour responses with or without the infusion of  
CC donor lymphocytes, and to activate or stimulate the immune system.  
CC The present sequence represents part of the human zalphal1 ligand gene  
CC containing the partial intron 3, exon 4, intron 4 and partial exon 5  
CC sequences.  
XX  
SQ Sequence 645 BP; 244 A; 107 C; 89 G; 204 T; 1 other;  
XX  
Query Match 92.0%; Score 18.4; DB 24; Length 645;  
Best Local Similarity 95.0%; Pred. No. 2.7e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AATGTTCTTAAGATTGTTT 20  
DB 383 AATGTTCTTAAGATTGTTT 364

RESULT 3  
ABL22726/c  
ID ABL22726 standard; DNA; 2696 BP.  
XX  
AC ABL22726;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 19651.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX  
OS Drosophila melanogaster.  
XX  
PN W0200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;



DR WPI; 2001-656860/75.  
 CC New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX Claim 1; SEQ ID NO 19651; 21bp + Sequence Listing; English.  
 PS  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB101840-AB116175) and the encoded proteins  
 CC sequences (AB101840-AB116175).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX Sequence 2696 BP; 919 A; 581 C; 512 G; 684 T; 0 other;  
 SQ  
 Query Match 92.0%; Score 18.4; DB 23; Length 2696;  
 Best Local Similarity 95.0%; Pred. No. 2.5e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AATGTTCTTAAGATTGTTT 20  
 Db 1704 AATGTTCTTAAGATTGTTT 1685  
 ID ABL27256 standard; DNA; 3273 BP.  
 XX ABL27256;  
 AC 26-MAR-2002 (first entry)  
 XX  
 DT Drosophila melanogaster genomic polynucleotide SEQ ID NO 33241.  
 XX  
 DE Drosophila; developmental biology; cell signalling; insecticide;  
 XX  
 KM pharmacological; gene; ds.  
 KM  
 XX Drosophila melanogaster.  
 OS  
 XX WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 XX  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EM;  
 XX  
 XX WPI; 2001-656860/75.  
 DR  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX Claim 1; SEQ ID NO 33241; 21bp + Sequence Listing; English.  
 PS  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA

CC sequences (AB101840-AB116175) and the encoded proteins  
 CC (AB101840-AB116175).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX Sequence 3273 BP; 1076 A; 658 C; 639 G; 900 T; 0 other;  
 SQ  
 Query Match 92.0%; Score 18.4; DB 23; Length 3273;  
 Best Local Similarity 95.0%; Pred. No. 2.5e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AATGTTCTTAAGATTGTTT 20  
 Db 776 AATGTTCTTAAGATTGTTT 795  
 ID ABL14642 standard; cDNA; 4839 BP.  
 XX ABL14642;  
 AC 26-MAR-2002 (first entry)  
 XX  
 DT Drosophila melanogaster expressed polynucleotide SEQ ID NO 38408.  
 XX  
 DE Drosophila; developmental biology; cell signalling; insecticide;  
 XX  
 KM pharmacological; gene; ss.  
 KM  
 XX Drosophila melanogaster.  
 OS  
 XX WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 XX  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EM;  
 XX  
 XX WPI; 2001-656860/75.  
 DR  
 XX P-PSDB; ABB70539.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX Claim 1; SEQ ID NO 38408; 21bp + Sequence Listing; English.  
 PS  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB101840-AB116175) and the encoded proteins  
 CC sequences (AB101840-AB116175).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX Sequence 4839 BP; 1360 A; 1051 C; 1092 G; 1336 T; 0 other;  
 SQ  
 Query Match 92.0%; Score 18.4; DB 23; Length 4839;  
 Best Local Similarity 95.0%; Pred. No. 2.5e+02;  
 Matches 19; Conservative 1; Indels 0; Gaps 0;  
 QY 1 AATGTTCTTAAGATTGTTT 20

DB 2454 AATGTTCTTAAGATTITTT 2435

|||||

RESULT 6  
ID ABL32039 standard; DNA; 6164 BP.  
XX ABL32039;  
AC ABL32039;  
XX 26-MAR-2002 (first entry)  
DT  
DE Human immune system associated gene SEQ ID NO: 12.  
XX  
KW Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antiangioma; cytosine; neoplastic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antirheumatic; antidiabetic; antipsoriatic;  
KW acute myeloid leukaemia; cancer; eye disease; arteriosclerosis; anaemia;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX WO200200928-A2.  
XX 03-JAN-2002.  
XX  
XX 02-JUL-2001; 2001WO-EP07537.  
XX  
XX 30-JUN-2000; 2000DE-1032529.  
XX 01-SEP-2000; 2000DE-1043826.  
XX  
XX (EPIG-) EPIGENOMICS AG.  
XX  
XX Olek A, Piepenbrock C, Berlin K;  
XX WPI; 2002-130909/17.  
XX  
XX Nucleic acid comprising fragment of chemically modified gene, useful  
XX for diagnosis and treatment of diseases associated with abnormal  
XX cytosine methylation  
XX  
XX  
XX Claim 1; SEQ ID NO 12; 32pp + Sequence listing; German.  
XX  
XX The present invention provides a number of human immune system associated  
XX genes which are modified by the methylation of cytosines. The sequences  
XX can be used in the diagnosis and treatment of immune system disorders,  
XX including eye diseases such as retinopathy, neovascular glaucoma and  
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
XX diseases. The present sequence is a gene of the invention.  
XX  
SQ Sequence 6164 BP; 1716 A; 134 C; 1458 G; 2855 T; 1 other;  
Query Match 92.0%; Score 18.4; DB 24; Length 6164;  
Best Local Similarity 95.0%; Pred. No. 2.5e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AATGTTCTTAAGATTITTT 20  
DB 4643 AATGTTCTTAAGATTITTT 4662

RESULT 7  
ID AAC29641 standard; CDNA; 149 BP.  
XX AAC29641;  
XX  
XX 06-OCT-2000 (first entry)  
DT

XX  
DE Human secreted protein 5' EST, SEQ ID NO: 33716.  
XX  
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX  
XX Homo sapiens.  
XX EP1033401-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 21-FEB-2000; 2000EP-0200610.  
XX  
XX 26-FEB-1999; 99US-0122487.  
XX  
XX (GSEST) GENSET.  
XX  
XX Dumas Malne Edwards J, Duclert A, Giordano J;  
XX WPI; 2000-500381/45.  
XX  
XX  
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX  
XX Claim 1; SEQ ID 33716; 71pp + CD-ROM; English.  
XX  
XX The present sequence is one of a large number of 5' ESTs derived from  
XX cDNAs encoding secreted proteins. No ORF has yet been conclusively  
XX identified within the present sequence. The 5' ESTs were prepared from  
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
XX sequences usually correspond mainly to the 3' untranslated region (UTR)  
XX of the mRNA because they are often obtained from oligo-dT primed cDNA  
XX libraries. Such ESTs are not well suited for isolating cDNA sequences  
XX derived from the 5' ends of mRNAs and even in those cases where longer  
XX cDNA sequences have been obtained, the full 5' UTR is rarely included.  
XX 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
XX used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
XX in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX They are used to obtain upstream regulatory sequences and to design  
XX expression and secretion vectors.  
XX  
SQ Sequence 149 BP; 39 A; 24 C; 25 G; 61 T; 0 other;  
Query Match 87.0%; Score 17.4; DB 21; Length 149;  
Best Local Similarity 94.7%; Pred. No. 6.9e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AATGTTCTTAAGATTITTT 19  
DB 37 AATGTTCTTAAGATTITTT 55

RESULT 8  
ID ABN63544/c standard; CDNA; 556 BP.  
XX ABN63544;  
XX  
XX 28-JUN-2002 (first entry)  
DT  
DE Human cancer related polynucleotide SEQ ID NO 3511.  
XX  
KW Human; cytosine; gene expression; gene mapping; tissue profiling;  
KW gene therapy; cancer; tumour; gene; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200214500-A2.  
XX  
XX 21-FEB-2002.  
XX

PF 16-AUG-2001; 2001MO-US25840.  
XX  
PR 16-AUG-2000; 2000US-226326P.  
XX  
PA (CHIR-) CHIRON CORP.  
PA (HXS-) HXSQ INC.  
XX  
PI Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F,  
PI Lamson G, Scott EM, Zhang G, Kaessam A, Pot D, Labat I,  
PI WPI; 2002-241905/29.  
DR  
XX  
PT New nucleic acid for producing a polypeptide, detecting differentially  
PT expressed genes correlated with a cancerous state of a mammalian cell,  
PT and inhibiting tumor growth -  
XX  
PS Claim 1; SEQ ID NO 3511; 883bp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated polynucleotide (ABN27253-ABN33262)  
CC with cytostatic activity. The polynucleotide is used to produce a  
CC polypeptide, to detect differentially expressed genes correlated with a  
CC cancerous state of a mammalian cell and to inhibit tumor growth. The  
CC polynucleotide is used as a probe in mapping and tissue profiling. The  
CC encoded polypeptide and antibodies to the polypeptide can also be used  
CC for therapeutic and diagnostic purposes. The polynucleotide is useful for  
CC gene therapy.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 556 BP; 186 A; 111 C; 120 G; 139 T; 0 other;  
XX  
Query Match 87.0%; Score 17.4; DB 24; Length 556;  
Best Local Similarity 94.7%; Pred. No. 6.6e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
XX  
QY 2 ATGTTCTTAAGATTCTTT 20  
DB 245 ATGTTCTTAAGATTCTTT 227  
XX  
RESULT 9  
ABN71515/c  
ID ABN71515 standard; DNA; 954 BP.  
XX  
AC ABN71515;  
XX  
DT 01-JUL-2002 (first entry)  
XX  
DE Streptococcus polynucleotide SEQ ID NO 10943.  
XX  
KM Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;  
KM group A streptococcus; Streptococcus pyogenes; antibacterial; gene;  
KM antiinflammatory; infection; vaccine; meningitis; gene therapy; de.  
XX  
OS Streptococcus agalactiae.  
XX  
PN WO200234771-A2.  
XX  
PD 02-MAY-2002.  
XX  
PF 29-OCT-2001; 2001MO-GB04789.  
XX  
PR 27-OCT-2000; 2000GB-0026333.  
PR 24-NOV-2000; 2000GB-0028727.  
PR 07-MAR-2001; 2001GB-0005640.  
XX  
PA (CHIR-) CHIRON SPA.  
PA (GENO-) INST GENOMIC RES.  
XX  
PI Telford J, Maignani V, Margarit Ros YI, Grandi G, Frazer C,  
PI Tettelein H;  
XX

DR WPI; 2002-352536/38.  
DR P-PSDB; ABP30884.  
XX  
XX New Streptococcus protein for the treatment or prevention of infection  
PT or disease caused by Streptococcus bacteria, such as meningitis, and  
PT for detecting a compound that binds to the protein -  
XX  
PS Claim 7; Page 4191; 4525bp; English.  
XX  
XX The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (1), nucleic acids encoding (1), ABN6044-ABN71526 and  
CC antibodies that bind (1) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (1) are used to detect Streptococcus in a  
CC biological sample. (1) is used to determine whether a compound binds to  
CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (1) may be used to recombinantly produce (1) and may be  
CC used in gene therapy. Antibodies to (1) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins.  
XX  
SQ Sequence 954 BP; 293 A; 145 C; 217 G; 299 T; 0 other;  
XX  
Query Match 87.0%; Score 17.4; DB 24; Length 954;  
Best Local Similarity 94.7%; Pred. No. 6.5e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
XX  
QY 2 ATGTTCTTAAGATTCTTT 20  
DB 764 ATGTTCTTAAGATTCTTT 746  
XX  
RESULT 10  
ABN66487/c  
ID ABN66487 standard; DNA; 996 BP.  
XX  
AC ABN66487;  
XX  
DT 01-JUL-2002 (first entry)  
XX  
DE Streptococcus polynucleotide SEQ ID NO 887.  
XX  
KM Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;  
KM group A streptococcus; Streptococcus pyogenes; antibacterial; gene;  
KM antiinflammatory; infection; vaccine; meningitis; gene therapy; de.  
XX  
OS Streptococcus agalactiae.  
XX  
PN WO200234771-A2.  
XX  
PD 02-MAY-2002.  
XX  
PF 29-OCT-2001; 2001MO-GB04789.  
XX  
PR 27-OCT-2000; 2000GB-0026333.  
PR 24-NOV-2000; 2000GB-0028727.  
PR 07-MAR-2001; 2001GB-0005640.  
XX  
PA (CHIR-) CHIRON SPA.  
PA (GENO-) INST GENOMIC RES.  
XX  
PI Telford J, Maignani V, Margarit Ros YI, Grandi G, Frazer C,  
PI Tettelein H;  
XX  
DR WPI; 2002-352536/38.  
DR P-PSDB; ABP25856.  
XX

PT New Streptococcus protein for the treatment or prevention of infection  
 PT or disease caused by Streptococcus bacteria, such as meningitis, and  
 PT for detecting a compound that binds to the protein -

PS Claim 7; Page 3241-3242; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and  
 CC antibodies that bind (1) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a  
 CC biological sample. (1) is used to determine whether a compound binds to  
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (1) may be used to recombinantly produce (1) and may be  
 CC used in gene therapy. Antibodies to (1) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins.

SO Sequence 996 BP; 309 A; 148 C; 229 G; 310 T; 0 other;

Query Match 87.0%; Score 17.4; DB 24; Length 996;  
 Best Local Similarity 94.7%; Pred. No. 6.5e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ATGTTCTTAAGATTGTTT 20  
 ||| ||||| ||||| |||||

Db 806 ATGGTCTTAAGATTGTTT 788

#### RESULT 11

ABN66079/C  
 ID ABN66079 standard; DNA; 1068 BP.

AC ABN66079;

DT 01-JUL-2002 (first entry)

XX Streptococcus polymyclotide SEQ ID NO 71.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

KM group A streptococcus; Streptococcus pyogenes; antibacterial; gene;  
 KM antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.

OS Streptococcus pyogenes.

XX WO200234771-A2.

PD 02-MAY-2002.

PF 29-OCT-2001; 2001WO-GB04789.

XX 27-OCT-2000; 2000GB-0026333.

PR 24-NOV-2000; 2000GB-0028727.

PR 07-MAR-2001; 2001GB-0005640.

PA (CHIR-) CHIRON SPA.

PI (GENO-) INST GENOMIC RES.  
 PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;  
 PI Tettelein H;

DR WPI; 2002-352536/38.

DR P-PSDB; ABP25448.

PT New Streptococcus protein for the treatment or prevention of infection  
 PT or disease caused by Streptococcus bacteria, such as meningitis, and  
 PT for detecting a compound that binds to the protein -

XX Claim 7; Page 3161; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and  
 CC antibodies that bind (1) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a  
 CC biological sample. (1) is used to determine whether a compound binds to  
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (1) may be used to recombinantly produce (1) and may be  
 CC used in gene therapy. Antibodies to (1) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins.

SO Sequence 1068 BP; 341 A; 204 C; 224 G; 299 T; 0 other;

Query Match 87.0%; Score 17.4; DB 24; Length 1068;  
 Best Local Similarity 94.7%; Pred. No. 6.4e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ATGTTCTTAAGATTGTTT 20  
 ||| ||||| ||||| |||||

Db 1043 ATTTTCTTAAGATTGTTT 1025

#### RESULT 12

AAK39670  
 ID AAK39670 standard; DNA; 1528 BP.

AC AAK39670;

DT 02-JUL-1999 (first entry)

XX Renal cancer associated gene.

KM Cancer associated antigen; diagnosis; research; treatment; human;  
 KM breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;

KM prostate cancer; ss.

OS Homo sapiens.

PN WO9904265-A2.

XX 28-JAN-1999.

PF 15-JUL-1998; 98WO-US14679.

XX 22-JUN-1998; 98US-0102322.

PR 17-JUL-1997; 97US-0896164.

PR 10-OCT-1997; 97US-0061599.

PR 10-OCT-1997; 97US-0061765.

PR 10-OCT-1997; 97US-0948705.

PR 11-OCT-1997; 97GB-0021697.

PA (LUDW-) LUDWIG INST CANCER RES.

PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;  
 PI Pfeundschuh M, Sahin U, Scanlan MJ, Stockert E;  
 PI Tureci O;

DR WPI; 1999-132448/11.

PT New isolated cancer associated nucleic acids and polypeptides -  
 PT isolated using sera from cancer patients, used to develop products  
 PT for the diagnosis, monitoring or treatment of cancers



DR WPI, 2001-656860/75.  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 XX Claim 1, SEQ ID NO 1981, 21pp + Sequence Listing, English.  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins  
 CC (AB57737-AB572072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 CC Sequence 2483 BP; 787 A; 474 C; 454 G; 768 T; 0 other;  
 SQ  
 Query Match 87.0%; Score 17.4; DB 23; Length 2483;  
 Best Local Similarity 94.7%; Pred. No. 6.2e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 2 AATGTTCTTAAGATTTT 20  
 2339 AAGTTCTTAAGATTTT 2357  
 DB  
 RESULT 15  
 ABL1262  
 ID ABL1262 standard; cDNA; 2531 BP.  
 XX  
 AC ABL1262;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE *Drosophila melanogaster* expressed polynucleotide SEQ ID NO 28268.  
 XX  
 XX *Drosophila*; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 XX  
 OS *Drosophila melanogaster*.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PT Venter JC, Adams M, Li PWD, Myers EW;  
 PT WPI, 2001-656860/75.  
 DR P-PSDB; ABB67159.  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 XX Claim 1, SEQ ID NO 28268; 21pp + Sequence Listing, English.  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins  
 CC (AB57737-AB572072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 CC Sequence 2531 BP; 748 A; 489 C; 587 G; 707 T; 0 other;  
 SQ  
 Query Match 87.0%; Score 17.4; DB 23; Length 2531;  
 Best Local Similarity 94.7%; Pred. No. 6.2e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 AATGTTCTTAAGATTTT 19  
 2253 AATGTTCTTAAGATTTT 2271  
 DB  
 Search completed: August 1, 2003, 13:55:23  
 Job time : 8.69866 secs

Sun Aug 3 09:04:02 2003

us-09-834-291-15.rnpb

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:05:17 ; Search time 5.18607 Seconds  
(without alignments)  
7955.924 Million cell updates/sec

Title: US-09-834-291-15

Perfect score: 1 aatgcttctaagattttt 20

Sequence: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Scoring table: 1439767 seqs, 1031500376 residues

Searched: 2879534

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
4: /cn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
5: /cn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
6: /cn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
7: /cn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
8: /cn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
9: /cn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
10: /cn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
11: /cn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
12: /cn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
13: /cn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
14: /cn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
16: /cn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*  
17: /cn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	US-09-834-291-15	Sequence 15, Appl
2	18.4	92.0	645	US-09-923-246-108	Sequence 108, App
3	18.4	92.0	645	US-10-295-723-108	Sequence 224463, A
4	18	90.0	587	US-10-027-632-224463	Sequence 21892, A
5	17.4	87.0	733	US-10-027-632-21892	Sequence 14, Appl
6	16.8	84.0	20	US-09-834-291-14	Sequence 20, Appl
7	16.8	84.0	40	US-09-834-291-20	Sequence 28, Appl
8	16.8	84.0	40	US-09-834-291-28	Sequence 455, App
9	16.8	84.0	366	US-09-764-877-455	Sequence 13025, A
10	16.8	84.0	450	US-09-960-352-13025	Sequence 21175, A
11	16.8	84.0	479	US-09-918-995-21175	Sequence 104, App
12	16.8	84.0	515	US-10-027-632-245649	Sequence 245649, A
13	16.8	84.0	578	US-10-027-632-229994	Sequence 229994, A
14	16.8	84.0	583	US-10-027-632-239669	Sequence 239669, A
15	16.8	84.0	603	US-10-027-632-239669	Sequence 200121, A
16	16.8	84.0	650	US-10-027-632-200121	Sequence 200121, A

17	16.8	84.0	666	US-10-027-632-256867	Sequence 256867, A
18	16.8	84.0	666	US-10-027-632-256868	Sequence 256868, A
19	16.8	84.0	666	US-10-027-632-256870	Sequence 256870, A
20	16.8	84.0	666	US-10-027-632-256871	Sequence 256871, A
21	16.8	84.0	666	US-10-027-632-170398	Sequence 170398, A
22	16.8	84.0	736	US-10-027-632-121985	Sequence 121985, A
23	16.8	84.0	980	US-10-027-632-9152	Sequence 9152, App
24	16.8	84.0	993	US-09-815-242-7222	Sequence 7222, App
25	16.8	84.0	1224	US-10-027-632-203488	Sequence 203488, A
26	16.8	84.0	1843	US-09-938-842A-3068	Sequence 3068, App
27	16.8	84.0	2272	US-10-114-170-126	Sequence 126, App
28	16.8	84.0	2827	US-09-834-291-4	Sequence 4, Appl
29	16.8	84.0	3212	US-09-834-291-1	Sequence 1, Appl
30	16.8	84.0	3212	US-09-967-552A-72	Sequence 72, Appl
31	16.8	84.0	27483	US-09-764-877-2928	Sequence 2928, App
32	16.8	84.0	640681	US-09-790-988-1	Sequence 1, Appl
33	16.8	84.0	169139	US-10-067-514-1	Sequence 1, Appl
34	16.8	84.0	336	US-09-834-975-358	Sequence 358, App
35	16.4	82.0	412	US-09-918-995-17530	Sequence 17530, A
36	16.4	82.0	510	US-09-833-381-1375	Sequence 1375, App
37	16.4	82.0	549	US-10-066-542A-2878	Sequence 2878, App
38	16.4	82.0	571	US-10-027-632-193079	Sequence 193079, A
39	16.4	82.0	602	US-10-027-632-13561	Sequence 13561, A
40	16.4	82.0	602	US-10-027-632-13562	Sequence 13562, A
41	16.4	82.0	617	US-10-027-632-304588	Sequence 304588, A
42	16.4	82.0	660	US-10-027-632-255212	Sequence 255212, A
43	16.4	82.0	684	US-10-027-632-84724	Sequence 84724, A
44	16.4	82.0	2000	US-09-938-842A-4583	Sequence 4583, App
45	16.4	82.0	2000	US-09-938-842A-4583	Sequence 4583, App

#### ALIGNMENTS

RESULT 1  
US-09-834-291-15  
; Sequence 15, Application US/09834291  
; Patent No. US20020042064A1  
; GENERAL INFORMATION:  
; APPLICANT: Krammer, Peter  
; APPLICANT: Muller-Schilling, Martina  
; APPLICANT: Oren, Moshe  
; TITLE OF INVENTION: p53 Binding Areas  
; FILE REFERENCE: 4121-122  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/DE99/03343  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: DE 198 47 779.1  
; PRIOR FILING DATE: 1998-10-16  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-09-834-291-15

Query Match 100.0%; Score 20; DB 9; Length 20;  
Best local Similarity 100.0%; Pred. No. 43;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 AATGTTCTTAAGATTGTTT 20  
Db 1 AATGTTCTTAAGATTGTTT 20

RESULT 2  
US-09-923-246-108/c  
; Sequence 108, Application US/09923246  
; Patent No. US20020128446A1  
; GENERAL INFORMATION:  
; APPLICANT: No. US20020128446A1ak, Julia E.

```
APPLICANT: Preenell, Scott R.
APPLICANT: Sprecher, Cindy A.
APPLICANT: Foster, Donald C.
APPLICANT: Holly, Richard D.
APPLICANT: Gross, Jane A.
APPLICANT: Johnston, Janet V.
APPLICANT: Nelson, Andrew J.
APPLICANT: Dillon, Stacey R.
APPLICANT: Hammond, Angela K.
TITLE OF INVENTION: NOVEL CYTOKINE ZALPHA11 LIGAND
CURRENT APPLICATION NUMBER: US/09/923,246
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/522,217
PRIOR FILING DATE: EARLIER FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/123,904
PRIOR FILING DATE: EARLIER FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/142,013
NUMBER OF SEQ ID NOS: 115
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 108
LENGTH: 645
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(645)
OTHER INFORMATION: n = A,T,C or G
US-09-923-246-108
```

## Query Match

Best Local Similarity 92.0%; Score 18.4; DB 10; Length 645;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 AATGTTCTTAAGATTTT 20
Db 383 AATGTTCTTAAGATTTT 364
```

```
RESULT 3
US-10-295-723-108/c
Sequence 108, Application US/10295723
Publication No. US20030125524A1
GENERAL INFORMATION:
APPLICANT: No. US20030125524A1ak, Julia E.
APPLICANT: Preenell, Scott R.
APPLICANT: Sprecher, Cindy A.
APPLICANT: Foster, Donald C.
APPLICANT: Holly, Richard D.
APPLICANT: Gross, Jane A.
APPLICANT: Johnston, Janet V.
APPLICANT: Nelson, Andrew J.
APPLICANT: Dillon, Stacey R.
APPLICANT: Hammond, Angela K.
TITLE OF INVENTION: NOVEL CYTOKINE ZALPHA11 LIGAND
FILE REFERENCE: 99-16
CURRENT APPLICATION NUMBER: US/10/295,723
PRIOR FILING DATE: 2002-11-15
PRIOR APPLICATION NUMBER: 09/522,217
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: US 60/123,547
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: US 60/123,904
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: US 60/142,013
NUMBER OF SEQ ID NOS: 115
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 108
LENGTH: 645
TYPE: DNA
ORGANISM: Homo sapiens
```

```
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(645)
OTHER INFORMATION: n = A,T,C or G
US-10-295-723-108
```

## Query Match

Best Local Similarity 92.0%; Score 18.4; DB 14; Length 645;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 AATGTTCTTAAGATTTT 20
Db 383 AATGTTCTTAAGATTTT 364
```

```
RESULT 4
US-10-027-632-224463
Sequence 224463, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: 2000-07-12
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 224463
LENGTH: 587
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(587)
OTHER INFORMATION: n = A,T,C or G
US-10-027-632-224463
```

## Query Match

Best Local Similarity 90.0%; Score 18; DB 13; Length 587;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 2 AATGTTCTTAAGATTTT 19
Db 134 AATGTTCTTAAGATTTT 151
```

```
RESULT 5
US-10-027-632-21892/c
Sequence 21892, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
```



PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 21892  
LENGTH: 733  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-21892

Query Match 87.0%; Score 17.4; DB 13;  
Best Local Similarity 94.7%; Pred. No. 7.3e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTGTTT 19  
DB 253 AATGTTCTTAAGATTGTTT 235

RESULT 6  
US-09-834-291-14  
Sequence 14, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 14  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-14

Query Match 84.0%; Score 16.8; DB 9;  
Best Local Similarity 90.0%; Pred. No. 8.1e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTGTTT 20  
DB 1 AATGTTCTTAAGATTGTTT 20

RESULT 7  
US-09-834-291-20  
Sequence 20, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21

PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 20  
LENGTH: 40  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-20

Query Match 84.0%; Score 16.8; DB 9;  
Best Local Similarity 90.0%; Pred. No. 8.8e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTGTTT 20  
DB 11 AATGTTCTTAAGATTGTTT 30

RESULT 8  
US-09-834-291-28  
Sequence 28, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 28  
LENGTH: 40  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-28

Query Match 84.0%; Score 16.8; DB 9;  
Best Local Similarity 90.0%; Pred. No. 8.8e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTGTTT 20  
DB 11 AATGTTCTTAAGATTGTTT 30

RESULT 9  
US-09-764-877-455/c  
Sequence 455, Application US/09764877  
Patent No. US20020147140A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC005  
CURRENT APPLICATION NUMBER: US/09/764,877  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - refer to PALM or file wrapper  
NUMBER OF SEQ ID NOS: 4031  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 455  
LENGTH: 366  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE

LOCATION: (302)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (310)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (313)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-877-455

Query Match  
Best Local Similarity 84.0%; Score 16.8; DB 10; Length 366;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTGTTT 20  
DB 274 AATGTTCTTAAGATTGTTT 255

RESULT 10  
US-09-960-352-13025  
Sequence 13025, Application US/09960352  
Patent No. US20020137139A1  
GENERAL INFORMATION:  
APPLICANT: Warren, Wesley C.  
APPLICANT: Tao, Nengbing  
APPLICANT: Byate, John C.  
TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION  
FILE REFERENCE: 16511.006/37-21(10298)C  
CURRENT APPLICATION NUMBER: US/09/960,352  
CURRENT FILING DATE: 2001-09-24  
NUMBER OF SEQ ID NOS: 15112  
SEQ ID NO 13025  
LENGTH: 450  
TYPE: DNA  
ORGANISM: Bos taurus  
OTHER INFORMATION: Clone ID: 56-BOWMS1-023-Q1-E1-F8  
US-09-960-352-13025

Query Match  
Best Local Similarity 84.0%; Score 16.8; DB 10; Length 450;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTGTTT 20  
DB 127 AATGTTCTTAAGATTGTTT 146

RESULT 11  
US-09-918-995-21175/C  
Sequence 21175, Application US/09918995  
Patent No. US20030073623A1  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc.  
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
FILE REFERENCE: 20411-756  
CURRENT APPLICATION NUMBER: US/09/918,995  
CURRENT FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: US/09/235,076  
PRIOR FILING DATE: 1999-01-20  
NUMBER OF SEQ ID NOS: 38054  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 21175  
LENGTH: 479  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)-(479)  
OTHER INFORMATION: n = A,T,C or G

US-09-918-995-21175

Query Match  
Best Local Similarity 84.0%; Score 16.8; DB 11; Length 479;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTGTTT 20  
DB 152 AATGTTCTTAAGATTGTTT 133

RESULT 12  
US-09-811-284-104/C  
Sequence 104, Application US/09811284  
Patent No. US20020058306A1  
GENERAL INFORMATION:  
APPLICANT: Vogel, Gabriel  
TITLE OF INVENTION: No. US20020058306A1 G Protein-Coupled Receptors  
FILE REFERENCE: 00167US1  
CURRENT APPLICATION NUMBER: US/09/811,284  
CURRENT FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 60/189,783  
PRIOR FILING DATE: 2000-03-16  
PRIOR APPLICATION NUMBER: 60/189,907  
PRIOR FILING DATE: 2000-03-16  
PRIOR APPLICATION NUMBER: 60/189,918  
PRIOR FILING DATE: 2000-03-16  
PRIOR APPLICATION NUMBER: 60/189,960  
PRIOR FILING DATE: 2000-03-16  
PRIOR APPLICATION NUMBER: 60/189,917  
PRIOR FILING DATE: 2000-03-16  
PRIOR APPLICATION NUMBER: 60/192,945  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/192,916  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/192,923  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/192,933  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/192,830  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/192,234  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: 60/192,155  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: 60/192,935  
PRIOR FILING DATE: 2000-03-29  
SOFTWARE: PatentIn version 3.0  
NUMBER OF SEQ ID NOS: 258  
SEQ ID NO 104  
LENGTH: 515  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-811-284-104

Query Match  
Best Local Similarity 84.0%; Score 16.8; DB 9; Length 515;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTGTTT 20  
DB 299 AATGTTCTTAAGATTGTTT 280

RESULT 13  
US-10-027-632-245649/C  
Sequence 245649, Application US/10027632  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30  
 ; PRIOR APPLICATION NUMBER: US 60/218,006  
 ; PRIOR FILING DATE: 2000-07-12  
 ; PRIOR APPLICATION NUMBER: US 60/198,676  
 ; PRIOR FILING DATE: 2000-04-20  
 ; PRIOR APPLICATION NUMBER: US 60/193,483  
 ; PRIOR FILING DATE: 2000-03-29  
 ; PRIOR APPLICATION NUMBER: US 60/185,218  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: US 60/167,363  
 ; PRIOR FILING DATE: 1999-11-23  
 ; PRIOR APPLICATION NUMBER: US 60/156,358  
 ; PRIOR FILING DATE: 1999-09-28  
 ; PRIOR APPLICATION NUMBER: US 60/146,002  
 ; PRIOR FILING DATE: 1999-08-09  
 ; NUMBER OF SEQ ID NOS: 325720  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 245649  
 ; LENGTH: 578  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-10-027-632-245649

Query Match 84.0%; Score 16.8; DB 13; Length 578;  
 Best Local Similarity 90.0%; Pred. No. 1.2e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20  
 DB 499 AATGTTCTTAAGATTTT 480

RESULT 14  
 US-10-027-632-229994  
 ; Sequence 229994, Application US/10027632  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, David G.  
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
 ; FILE REFERENCE: 108827.129  
 ; CURRENT APPLICATION NUMBER: US/10/027,632  
 ; CURRENT FILING DATE: 2002-04-30  
 ; PRIOR APPLICATION NUMBER: US 60/218,006  
 ; PRIOR FILING DATE: 2000-07-12  
 ; PRIOR APPLICATION NUMBER: US 60/198,676  
 ; PRIOR FILING DATE: 2000-04-20  
 ; PRIOR APPLICATION NUMBER: US 60/193,483  
 ; PRIOR FILING DATE: 2000-03-29  
 ; PRIOR APPLICATION NUMBER: US 60/185,218  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: US 60/167,363  
 ; PRIOR FILING DATE: 1999-11-23  
 ; PRIOR APPLICATION NUMBER: US 60/156,358  
 ; PRIOR FILING DATE: 1999-09-28  
 ; PRIOR APPLICATION NUMBER: US 60/146,002  
 ; PRIOR FILING DATE: 1999-08-09  
 ; NUMBER OF SEQ ID NOS: 325720  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 229994  
 ; LENGTH: 583  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-10-027-632-229994

Query Match 84.0%; Score 16.8; DB 13; Length 583;  
 Best Local Similarity 90.0%; Pred. No. 1.2e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20  
 DB 123 AATGTTCTTAAGATTTT 142

RESULT 15  
 US-10-027-632-239669/c  
 ; Sequence 239669, Application US/10027632  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, David G.  
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
 ; FILE REFERENCE: 108827.129  
 ; CURRENT APPLICATION NUMBER: US/10/027,632  
 ; CURRENT FILING DATE: 2002-04-30  
 ; PRIOR APPLICATION NUMBER: US 60/218,006  
 ; PRIOR FILING DATE: 2000-07-12  
 ; PRIOR APPLICATION NUMBER: US 60/198,676  
 ; PRIOR FILING DATE: 2000-04-20  
 ; PRIOR APPLICATION NUMBER: US 60/193,483  
 ; PRIOR FILING DATE: 2000-03-29  
 ; PRIOR APPLICATION NUMBER: US 60/185,218  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: US 60/167,363  
 ; PRIOR FILING DATE: 1999-11-23  
 ; PRIOR APPLICATION NUMBER: US 60/156,358  
 ; PRIOR FILING DATE: 1999-09-28  
 ; PRIOR APPLICATION NUMBER: US 60/146,002  
 ; PRIOR FILING DATE: 1999-08-09  
 ; NUMBER OF SEQ ID NOS: 325720  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 239669  
 ; LENGTH: 603  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(603)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-10-027-632-239669

Query Match 84.0%; Score 16.8; DB 13; Length 603;  
 Best Local Similarity 90.0%; Pred. No. 1.2e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 20  
 DB 262 AATGTTCTTAAGATTTT 243

Search completed: August 1, 2003, 13:37:02  
 Job time : 6.18607 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 02:21:37 ; Search time 1.65066 Seconds  
(without alignments)  
5347.959 Million cell updates/sec

Title: US-09-834-291-15

Perfect score: 20  
Sequence: 1 aatgttcttaagaattttt 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/prodata/1/ina/PCITUS\_COMB.seq: \*  
6: /cgn2\_6/prodata/1/ina/backfileseq1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18.4	92.0	645	4 US-09-522-217-108	Sequence 108, App
C 2	16.8	84.0	2372	4 US-09-453-702B-116	Sequence 126, App
C 3	16.4	82.0	1770	3 US-08-943-731-146	Sequence 146, App
C 4	16.4	82.0	2388	4 US-09-205-258-154	Sequence 154, App
C 5	16.4	82.0	24183	3 US-08-943-731-3	Sequence 3, App1
C 6	16.4	82.0	1664976	4 US-08-916-421B-1	Sequence 1, App1
C 7	15.8	79.0	672	4 US-09-134-001C-1361	Sequence 1361, App
C 8	15.8	79.0	1619	1 US-08-507-455-1	Sequence 1, App1
C 9	15.8	79.0	1644	4 US-08-507-455-4	Sequence 4, App1
C 10	15.8	79.0	1754	1 US-08-507-455-2	Sequence 2, App1
C 11	15.8	79.0	3084	4 US-09-620-312D-616	Sequence 616, App
C 12	15.8	79.0	48974	3 US-08-920-422-17	Sequence 17, App
C 13	15.8	79.0	580073	4 US-08-545-528D-1	Sequence 1, App1
C 14	15.4	77.0	339	4 US-09-107-532A-13414	Sequence 13414, App
C 15	15.4	77.0	1020	4 US-09-107-532A-1250	Sequence 1250, App
C 16	15.4	77.0	2893	2 US-08-415-593-42	Sequence 42, App1
C 17	15.4	77.0	6718	2 US-08-962-284-1	Sequence 1, App1
C 18	15.4	77.0	7042	3 US-09-092-508-1	Sequence 1, App1
C 19	15.4	77.0	7042	4 US-09-435-115-1	Sequence 1, App1
C 20	15.4	77.0	7042	4 US-09-098-310-1	Sequence 1, App1
C 21	15.4	77.0	7042	4 US-09-092-364-21	Sequence 21, App1
C 22	15.4	77.0	7075	3 US-09-092-508-15	Sequence 15, App1
C 23	15.4	77.0	7075	4 US-09-435-115-15	Sequence 15, App1
C 24	15.2	76.0	106	2 US-08-184-009-165	Sequence 165, App
C 25	15.2	76.0	106	2 US-08-458-356-165	Sequence 165, App
C 26	15.2	76.0	106	2 US-08-460-736-165	Sequence 165, App
C 27	15.2	76.0	106	4 US-09-535-370-165	Sequence 165, App

C 28	15.2	76.0	264	2 US-08-607-384A-7	Sequence 7, App1
C 29	15.2	76.0	363	4 US-09-107-532A-1913	Sequence 1913, App
C 30	15.2	76.0	455	6 5278286-1	Patent No. 5278286
C 31	15.2	76.0	457	4 US-09-702-705-861	Sequence 861, App
C 32	15.2	76.0	457	4 US-09-736-457-861	Sequence 861, App
C 33	15.2	76.0	528	4 US-09-702-705-957	Sequence 957, App
C 34	15.2	76.0	528	4 US-09-736-457-957	Sequence 957, App
C 35	15.2	76.0	541	4 US-09-679-409-37	Sequence 37, App1
C 36	15.2	76.0	601	2 US-08-184-009-168	Sequence 168, App
C 37	15.2	76.0	601	2 US-08-458-356-168	Sequence 168, App
C 38	15.2	76.0	601	3 US-08-460-736-168	Sequence 168, App
C 39	15.2	76.0	601	4 US-09-535-370-168	Sequence 168, App
C 40	15.2	76.0	687	4 US-09-328-352-1113	Sequence 1113, App
C 41	15.2	76.0	705	4 US-09-107-532A-2425	Sequence 2425, App
C 42	15.2	76.0	789	4 US-09-134-001C-2695	Sequence 2695, App
C 43	15.2	76.0	813	3 US-09-028-819-13	Sequence 13, App1
C 44	15.2	76.0	813	4 US-09-556-163-13	Sequence 13, App1
C 45	15.2	76.0	915	4 US-09-601-198-90	Sequence 90, App1

## ALIGNMENTS

RESULT 1  
US-09-522-217-108/c  
Sequence 108, Application US/09522217  
Patent No. 6307024

GENERAL INFORMATION:

APPLICANT: NO. 6307024ak, Julia E.

APPLICANT: Presnell, Scott R.

APPLICANT: Spirecher, Cindy A.

APPLICANT: Foster, Donald C.

APPLICANT: Holly, Richard D.

APPLICANT: Gross, Jane A.

APPLICANT: Johnston, Janet V.

APPLICANT: Nelson, Andrew J.

APPLICANT: Dillon, Stacey R.

APPLICANT: Hammond, Angela K.

TITLE OF INVENTION: NOVEL CYTOKINE ZALPHALL LIGAND

FILE REFERENCE: 99-16

CURRENT APPLICATION NUMBER: US/09/522,217

CURRENT FILING DATE: 2000-03-09

EARLIER APPLICATION NUMBER: US 60/123,547

EARLIER FILING DATE: 1999-03-09

EARLIER APPLICATION NUMBER: US 60/123,904

EARLIER FILING DATE: 1999-03-11

EARLIER APPLICATION NUMBER: US 60/142,013

EARLIER FILING DATE: 1999-07-01

NUMBER OF SEQ ID NOS: 115

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 108

LENGTH: 645

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (1) ... (645)

OTHER INFORMATION: n = A,T,C or G

US-09-522-217-108

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 4; Length 645;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 AATGTTCTTAAGATTATTTT 20

383 AATGTTCTTAAGATTATTTT 364

RESULT 2  
US-09-453-702B-126/c  
Sequence 126, Application US/09453702B  
Patent No. 6365723

GENERAL INFORMATION:  
APPLICANT: Blattner, Frederick R.  
Burland, Valerie  
Perna, Nicole T.  
Plunkett, Guy  
Welch, Rod  
TITLE OF INVENTION: No. 6365723e1 Sequences of E. coli O157  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Quarles & Brady  
STREET: 1 South Pinckney Street  
CITY: Madison  
STATE: WI  
COUNTRY: US  
ZIP: 53701-2113  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 8.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/453,702B  
FILING DATE: 03-Dec-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/110,955  
FILING DATE: 04-DEC-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Seay, Nicholas J.  
REGISTRATION NUMBER: 27386  
REFERENCE/DOCKET NUMBER: 960296.95017  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 251-5000  
TELEFAX: (608) 251-9166  
INFORMATION FOR SEQ ID NO: 126:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2272  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 126:  
US-09-453-702B-126  
Query Match 84.0%; Score 16.8; DB 4; Length 2272;  
Best Local Similarity 90.0%; Pred. No. 1.5e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AATGTTCTTAAGATTGTTT 20  
DB 1176 AATGTGCTTATGATTGTTT 1157  
RESULT 3  
US-08-943-731-146  
; Sequence 146, Application US/08943731  
; Patent No. 6265157  
; GENERAL INFORMATION:  
; APPLICANT: PROCKOP, DARWIN J.  
; APPLICANT: SPOTILA, LORETTA D.  
; APPLICANT: DELTAS, CONSTANTINOS D.  
; APPLICANT: SEREDA, LARISA  
; APPLICANT: LARSON, ANDREA W.  
; APPLICANT: PACK, MICHAEL  
; APPLICANT: COLIGE, ALAIN  
; APPLICANT: EARLY, JAMES  
; APPLICANT: KORRKO, JARMO  
; APPLICANT: ALA-KORRKO, LEENA, et al.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING  
; TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES  
; NUMBER OF SEQUENCES: 666  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.

STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND  
STREET: FLR.  
CITY: PHILADELPHIA  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-7086  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/943,731  
FILING DATE: 03-OCT-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/212,322  
FILING DATE: 14-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/803,628  
FILING DATE: 03-DEC-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: DOYLE LEARY Ph.D., KATHRYN  
REGISTRATION NUMBER: 36,317  
REFERENCE/DOCKET NUMBER: 9598-27  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-965-1284  
TELEFAX: 215-567-2991  
TELEX: 831-494  
INFORMATION FOR SEQ ID NO: 146:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1770 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-943-731-146  
Query Match 82.0%; Score 16.4; DB 3; Length 1770;  
Best Local Similarity 94.4%; Pred. No. 2.2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3 TGTTCCTTAAGATTGTTT 20  
DB 44 TGTTCCTTAAGATTGTTAT 61  
RESULT 4  
US-09-205-258-154  
; Sequence 154, Application US/09205258  
; Patent No. 6525174  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 207 Human Secreted Proteins  
; FILE REFERENCE: P2007P1  
; CURRENT APPLICATION NUMBER: US/09/205,258  
; CURRENT FILING DATE: 1998-12-04  
; EARLIER APPLICATION NUMBER: PCT/US98/11422  
; EARLIER FILING DATE: 1998-06-04  
; EARLIER APPLICATION NUMBER: 60/048,885  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,375  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,881  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,880  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,896  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,020  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,876  
; EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,895  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,884  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,894  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,971  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,882  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,899  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,893  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,900  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,901  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,892  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,915  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,019  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,970  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,972  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,916  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,373  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,875  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,374  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,917  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,949  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,883  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,897  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,898  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,962  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30  
NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 154  
LENGTH: 2388  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-205-258-154  
Query Match 82.0%; Score 16.4; DB 4; length 2388;

Best Local Similarity 94.4%; Pred. No. 2.1e+02; Indels 0; Gaps 0;  
Matches 17; Conservative 0; Mismatches 1;  
QY 3 TGTTCCTTAAGATTTT 20  
Db 2334 TGTTCCTTAAGATTTT 2351  
RESULT 5  
US-08-943-731-3  
Sequence 3, Application US/08943731  
Patent No. 6265157  
GENERAL INFORMATION:  
APPLICANT: PROCKOP, DARWIN J.  
APPLICANT: SPOTILA, LORETTA D.  
APPLICANT: DELTA, CONSTANTINOS D.  
APPLICANT: SEREDA, LARISA  
APPLICANT: LARSON, ANDREA W.  
APPLICANT: PACK, MICHAEL  
APPLICANT: COLIGE, ALAIN  
APPLICANT: EARLY, JAMES  
APPLICANT: KORRKO, JARMO  
APPLICANT: ALA-KORRKO, LEENA, et al.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING  
TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES  
NUMBER OF SEQUENCES: 666  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P. C.  
STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND  
STREET: FLR.  
CITY: PHILADELPHIA  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-7086  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/943,731  
FILING DATE: 03-OCT-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/212,322  
FILING DATE: 14-MAR-1994  
PRIOR APPLICATION DATA: US 07/803,628  
APPLICATION NUMBER: US 07/803,628  
FILING DATE: 03-DEC-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: DOYLE LEARY Ph.D., KATHRYN  
REGISTRATION NUMBER: 36,317  
REFERENCE/DOCKET NUMBER: 9598-27  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-965-1284  
TELEFAX: 215-567-2991  
TELEX: 831-494  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 24183 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-943-731-3  
Query Match 82.0%; Score 16.4; DB 3; length 24183;  
Best Local Similarity 94.4%; Pred. No. 1.9e+02; Indels 0; Gaps 0;  
Matches 17; Conservative 1; Mismatches 1;  
QY 3 TGTTCCTTAAGATTTT 20  
Db 20352 TGTTCCTTAAGATTTTAT 20369

RESULT 6  
US-08-916-421B-1/c  
Sequence 1, Application US/08916421B  
Patent No. 6503729  
GENERAL INFORMATION:  
APPLICANT: Built et al.  
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus  
Patent No. 6503729  
FILE REFERENCE: PB275  
CURRENT APPLICATION NUMBER: US/08/916,421B  
PRIORITY FILING DATE: 1997-08-22  
PRIORITY FILING DATE: 1996-08-22  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 1664976  
TYPE: DNA  
ORGANISM: Methanococcus jannaschii  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (28222)..(28222)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (28257)..(28258)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (84773)..(84773)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (84808)..(84808)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (84812)..(84812)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98120)..(98120)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98159)..(98159)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98239)..(98239)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98266)..(98266)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98343)..(98343)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (103998)..(103998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (148548)..(148548)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (163385)..(163385)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (191989)..(191989)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (191995)..(191995)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (231980)..(231980)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234187)..(234187)  
OTHER INFORMATION: n equals a, t, c, or g

OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234220)..(234220)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234814)..(234814)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (309398)..(309398)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (309418)..(309418)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (312837)..(312837)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (312993)..(312993)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (319226)..(319226)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (559167)..(559167)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (559241)..(559241)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (600592)..(600592)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (622708)..(622708)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (657081)..(657081)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (657203)..(657203)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (674435)..(674435)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (682442)..(682442)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (713652)..(713652)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (741684)..(741684)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (779455)..(779455)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (779676)..(779676)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (855539)..(855539)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (871619)..(871619)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1084830)..(1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1096846)..(1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1119881)..(1119881)  
OTHER INFORMATION: n equals a, t, c, or g



NAME/KEY: misc\_feature  
LOCATION: (1130881)..(1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1349473)..(1349473)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1349491)..(1349491)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1470091)..(1470091)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1569020)..(1569020)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1602912)..(1602912)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1603734)..(1603734)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1637998)..(1637998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1664854)..(1664854)  
OTHER INFORMATION: n equals a, t, c, or g  
US-08-916-421B-1

Query Match 82.0%; Score 16.4; DB 4; Length 1664976;  
Best Local Similarity 94.4%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ATGTTCTTAAGATTTT 19  
Db 387346 ATGTTCTTAAGATTTT 387329

RESULT 7  
US-09-134-001C-1361  
Sequence 1361, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 1361  
LENGTH: 672  
TYPE: DNA  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-1361

Query Match 79.0%; Score 15.8; DB 4; Length 672;  
Best Local Similarity 89.5%; Pred. No. 3.9e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 19  
Db 647 AATGTTCTTAAGATTTT 665

RESULT 8  
US-08-507-455-1  
Sequence 1, Application US/08507455  
Patent No. 5695961  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: BIFUNCTIONAL EXPRESSION SYSTEM  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: USA  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/507,455  
FILING DATE: 08-SEP-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9303988.1  
FILING DATE: 26-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: MITCHARD, LEONARD C  
REGISTRATION NUMBER: 29,009  
REFERENCE/DOCKET NUMBER: 1498-72  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1619 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Saccharomyces cerevisiae  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: 546..547  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: 635..636  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: 1035..1036  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: 1411..1412  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 550..555  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 574..579  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 668..673  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 692..697  
US-08-507-455-1

Query Match 79.0%; Score 15.8; DB 1; Length 1619;  
Best Local Similarity 89.5%; Pred. No. 3.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AATGTTCTTAAGATTTT 19  
Db 1265 AATGATTATTAAGATTTT 1283

## RESULT 9

US-08-759-436-4/c  
Sequence 4, Application US/08759436  
Patent No. 6437218  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: DNA fragment encoding a protein  
TITLE OF INVENTION: Involved in  
TITLE OF INVENTION: fatty aldehyde decarboxylase activity, recombinant  
TITLE OF INVENTION: molecules comprising said fragment and a method for  
TITLE OF INVENTION: obtaining transformed bacterial cells and plants  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Iadad & Parry  
STREET: 26 West 61st Street  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10023  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/759,436  
FILING DATE: 5-Dec-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Cord, Janet I.  
REGISTRATION NUMBER: 33,778  
REFERENCE/DOCKET NUMBER: U011063-5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-708-1935  
TELEFAX: 212-246-8959  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1644 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Arabidopsis thaliana  
STRAIN: Landsberg erecta ecotype  
IMMEDIATE SOURCE:  
LIBRARY: Coates  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 1  
MAP POSITION: 12  
UNITS: CM  
PUBLICATION INFORMATION:  
AUTHORS: Aarts, Mark G.M.  
AUTHORS: Keijzer, Christiaan J.  
AUTHORS: Stiekema, Willem J.  
AUTHORS: Pereira, Andy  
TITLE: Molecular characterization of the CR1 gene  
TITLE: of Arabidopsis involved in epicuticular wax  
TITLE: biosynthesis and pollen fertility  
JOURNAL: Plant Cell  
VOLUME: 7  
ISSUE: 12  
PAGES: 2115-2127  
DATE: december-1995  
RELEVANT RESIDUES IN SEQ ID NO: 4: FROM 1 TO 750  
US-08-759-436-4

Query Match 79.0%; Score 15.8; DB 4; Length 1644;  
Best Local Similarity 89.5%; Pred. No. 3.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AATGTTCTTAAGATTTT 19  
Db 855 ACTGTTCTGAAGATTTT 837

## RESULT 10

US-08-507-455-2  
Sequence 2, Application US/08507455  
Patent No. 5695961  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: BIFUNCTIONAL EXPRESSION SYSTEM  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHUYE P.C.  
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: USA  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/507,455  
FILING DATE: 08-SEP-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9303988.1  
FILING DATE: 26-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: MITCHARD, LEONARD C  
REGISTRATION NUMBER: 29,009  
REFERENCE/DOCKET NUMBER: 1498-72  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1754 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Saccharomyces cerevisiae  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: 546..547  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: 635..636  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: 1035..1036  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: 1411..1412  
US-08-507-455-2

Query Match 79.0%; Score 15.8; DB 1; Length 1754;  
Best Local Similarity 89.5%; Pred. No. 3.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AATGTTCTTAAGATTTT 19  
Db 1397 AATGATTATTAAGATTTT 1415

## RESULT 11

US-09-620-312D-616  
Sequence 616, Application US/09620312D  
Patent No. 6569662

## GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Ren, Feiyun  
APPLICANT: Chen, Rui-hong  
APPLICANT: Zhao, Qing A.  
APPLICANT: Xu, Aigong J.  
APPLICANT: Yang, Yonghong  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yungang  
APPLICANT: Wang, Dunrui  
APPLICANT: Wang, Zhiwei  
APPLICANT: John Tillinghast  
APPLICANT: Drmanac, Radcoje T.  
TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
FILE REFERENCE: 784CIP2B  
CURRENT APPLICATION NUMBER: US/09/620,312D  
CURRENT FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 1105  
SOFTWARE: pc FL\_genes Version 1.0  
SEQ ID NO 616  
LENGTH: 3084  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (633)..(1883)  
US-09-620-312D-616

Query Match 79.0%; Score 15.8; DB 4; Length 3084;  
Best Local Similarity 89.5%; Pred. No. 3.6e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 19  
DB 3012 AAGTTCTTAAGATTTT 3030

## RESULT 12

US-08-920-422-17  
Sequence 17, Application US/08920422A  
Patent No. 6255473

## GENERAL INFORMATION:

APPLICANT: Vitek, Michael P.  
APPLICANT: Mltuda, No. 6255473iaki  
APPLICANT: Roses, Allen D.  
TITLE OF INVENTION: Presentin-1 Gene Promoter  
FILE REFERENCE: VITEKPRESENTIN  
CURRENT APPLICATION NUMBER: US/08/920,422A  
CURRENT FILING DATE: 1997-08-29  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 17  
LENGTH: 48974  
TYPE: DNA  
ORGANISM: Mus musculus  
US-08-920-422-17

Query Match 79.0%; Score 15.8; DB 3; Length 48974;  
Best Local Similarity 89.5%; Pred. No. 3e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 19  
DB 15475 AATTTTATTAAGATTTT 15493

## RESULT 13

US-08-545-528D-1  
Sequence 1, Application US/08545528D  
Patent No. 6537773

## GENERAL INFORMATION:

APPLICANT: Fraser et al.  
TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragments  
PATENT NO. 6537773  
FILE REFERENCE: PB193P1  
CURRENT APPLICATION NUMBER: US/08/545,528D  
CURRENT FILING DATE: 1995-10-19  
PRIOR APPLICATION NUMBER: US 08/488,018  
PRIOR FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: US 08/473,545  
PRIOR FILING DATE: 1995-06-07  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 1  
LENGTH: 580073  
TYPE: DNA  
ORGANISM: Mycoplasma genitalium  
US-08-545-528D-1

Query Match 79.0%; Score 15.8; DB 4; Length 580073;  
Best Local Similarity 89.5%; Pred. No. 2.2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTTCTTAAGATTTT 19  
DB 547054 AATGTTCTTAAGATTTT 547072

RESULT 14  
US-09-107-532A-3414  
Sequence 3414, Application US/09107532A  
Patent No. 6583275

## GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSER: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Walham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354

## COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Atinifello, Pamela Deneke

REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 3414:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 339 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...339  
SEQUENCE DESCRIPTION: SEQ ID NO: 3414:  
US-09-107-532A-3414

Query Match 77.0%; Score 15.4; DB 4; Length 339;  
Best Local Similarity 94.1%; Pred. No. 5.8e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAGATT 17  
|||||  
Db 321 AATGTTCTTAGATT 337

RESULT 15  
US-09-107-532A-1250/C  
Sequence 1250, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7110  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
City: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 1250:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1020 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...1020  
SEQUENCE DESCRIPTION: SEQ ID NO: 1250:  
US-09-107-532A-1250

Query Match 77.0%; Score 15.4; DB 4; Length 1020;  
Best Local Similarity 94.1%; Pred. No. 5.5e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGTTCTTAGATT 17  
|||||  
Db 201 AATGTTCTTAGATT 185

Search completed: August 1, 2003, 08:37:20  
Job time: 4.65066 secs

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Length	DB	ID	Description
1	20	100.0	20	6	AX026093
2	20	100.0	20	6	AX026098
3	20	100.0	20	6	AX026112
4	20	100.0	266	6	AX026120
5	20	100.0	266	6	HSN011034
6	20	100.0	720	6	AX026090
7	20	100.0	2380	6	AX026891
8	20	100.0	2827	6	AX026092
9	20	100.0	3312	6	AX026089
10	20	100.0	45121	6	AX695635
11	20	100.0	187313	6	AL157394
12	18.4	92.0	167412	2	AX026094
13	18.4	92.0	192657	2	AL156352
14	18.4	92.0	193352	10	AL672160
15	18.4	92.0	193352	2	BX294664
16	18.4	92.0	199776	9	AC011429
17	18.4	92.0	204012	10	AL1732613
18	90.0	81200	2	AC022086	
19	90.0	145667	10	AL662817	
20	90.0	214440	2	AC091422	
21	18	90.0	215973	8	AP444228
22	17.4	87.0	80961	5	AL139977
23	17.4	87.0	163933	5	AL192559
24	17.4	87.0	166772	9	AL1589763
25	17.4	87.0	175770	2	AC127943
26	17.4	87.0	185049	9	AL1365495
27	17.4	87.0	186325	9	AC074244
28	17.4	87.0	189485	2	AC141268
29	17.4	87.0	193347	2	AC125884
30	17.4	87.0	203246	5	AL1356606
31	17.4	87.0	206373	5	AL1954133
32	17.4	87.0	228659	2	AC115992
33	17.4	87.0	228659	2	AC116033
34	17.4	87.0	230097	2	AC126053
35	17.4	87.0	230097	2	AC099140
36	17.4	87.0	258131	2	AC135751
37	17.4	87.0	289818	2	AC108332
38	17.4	85.0	236723	2	AC112477
39	17	85.0	240214	2	AC126871
40	17	85.0	257842	2	AC105569
41	17	85.0	262681	2	AC128975
42	17	85.0	280469	2	AC121541
43	16.8	84.0	20	6	AX026095
44	16.8	84.0	20	6	AX026097

RESULT 1			
AX026093			
LOCUS	AX026093	20 bp	DNA
DEFINITION	Sequence 5 from Patent DE19847779.		linear
ACCESSION	AX026093		
VERSION	AX026093.1	GI:10187524	
KEYWORDS			
SOURCE	.		
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
	Eukaryota; Metazoa; Chordata; Carnathini; Homnidae; Homo.		

REFERENCE	1
AUTHORS	Mueller-Schilling, M., Krammer, P. and Oren, M.
TITLE	Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy

## JOURNAL

Patent: DE 19847779-C 5 03-FEB-2000;  
DEUTSCHES KREBSFORSCH (DE)

FEATURES  
SOURCE  
1. 20  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT  
ORIGIN  
7 a 7 c 5 g 1 t

Query Match  
Best Local Similarity 100.0%; Score 20; DB 6; Length 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  
1 GGACAGCCCTGACAGCCA 20  
1 GGACAGCCCTGACAGCCA 20

## RESULT 2

AX026098  
LOCUS  
DEFINITION Sequence 10 from Patent DE19847779.  
ACCESSION AX026098  
VERSION AX026098.1 GI:10187529  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Muller-Schilling, M., Krammer, P. and Oren, M.  
Novel receptor dna useful for identifying apoptosis-modulating  
substances potentially useful for cancer chemotherapy  
Patent: DE 19847779-C 10 03-FEB-2000;  
DEUTSCHES KREBSFORSCH (DE)

FEATURES  
SOURCE  
1. 20  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT  
ORIGIN  
7 a 7 c 5 g 1 t

Query Match  
Best Local Similarity 100.0%; Score 20; DB 6; Length 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  
1 GGACAGCCCTGACAGCCA 20  
1 GGACAGCCCTGACAGCCA 20

## RESULT 3

AX026112  
LOCUS  
DEFINITION Sequence 24 from Patent DE19847779.  
ACCESSION AX026112  
VERSION AX026112.1 GI:10187543  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Muller-Schilling, M., Krammer, P. and Oren, M.  
Novel receptor dna useful for identifying apoptosis-modulating  
substances potentially useful for cancer chemotherapy  
Patent: DE 19847779-C 24 03-FEB-2000;  
DEUTSCHES KREBSFORSCH (DE)

FEATURES  
SOURCE  
1. 20  
/organism="Homo sapiens"

BASE COUNT  
ORIGIN  
7 a 7 c 5 g 1 t

Query Match  
Best Local Similarity 100.0%; Score 20; DB 6; Length 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  
1 GGACAGCCCTGACAGCCA 20  
1 GGACAGCCCTGACAGCCA 20

## RESULT 4

AX026120  
LOCUS  
DEFINITION Sequence 32 from Patent DE19847779.  
ACCESSION AX026120  
VERSION AX026120.1 GI:10187551  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Muller-Schilling, M., Krammer, P. and Oren, M.  
Novel receptor dna useful for identifying apoptosis-modulating  
substances potentially useful for cancer chemotherapy  
Patent: DE 19847779-C 32 03-FEB-2000;  
DEUTSCHES KREBSFORSCH (DE)

FEATURES  
SOURCE  
1. 266  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT  
ORIGIN  
49 a 72 c 110 g 35 t

Query Match  
Best Local Similarity 100.0%; Score 20; DB 6; Length 266;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  
1 GGACAGCCCTGACAGCCA 20  
160 GGACAGCCCTGACAGCCA 179

## RESULT 5

HS011034  
LOCUS  
DEFINITION Homo sapiens DNA for enhancer of CD95 gene, partial.  
ACCESSION AV011034  
VERSION AV011034.1 GI:4165483  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Muller, M., Wilder, S., Bannasch, D., Israel, D., Lehlbach, K.,  
Krammer, P.H., Friedmann, S.L., Galle, P.R., Streimel, W., Oren, M. and  
p53 activates the CD95 (APO-1/Fas) gene in response to DNA damage  
by anticancer drugs  
J. Exp. Med. 188 (11), 2033-2045 (1998)

FEATURES  
SOURCE  
1. 266  
/organism="Homo sapiens"

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (08-SEP-1998) Muller M., Department of Internal Medicine  
IV, University Hospital, Bergheimerstr. 58, Heidelberg, 69115,

GERMANY  
Location/Qualifiers  
1. .266  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

gene  
/gene="CD95"  
159. .178  
/gene="CD95"  
/evidence="experimental"

enhancer  
49 a 72 c 110 g 35 t

BASE COUNT  
ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 266;  
100.0%; Pred. No. 40;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  
1 GGACAGCCCTGACAGCCA 20  
|||||  
160 GGACAGCCCTGACAGCCA 179

RESULT 6  
LOCUS AX026090 720 bp DNA linear PAT 16-SEP-2000  
DEFINITION Sequence 2 from Patent DE19847779.  
ACCESSION AX026090  
VERSION AX026090.1 GI:10187521

KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Mueller-Schilling, M., Kramer, P. and Oren, M.  
TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy  
JOURNAL Patent: DE 19847779-C 2 03-FEB-2000;  
DEUTSCHES KREBSFORSCH (DE)

FEATURES  
source location/Qualifiers  
1. .720  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 142 a 181 c 216 g 181 t

ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 20; DB 6; Length 720;  
100.0%; Pred. No. 33;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  
1 GGACAGCCCTGACAGCCA 20  
|||||  
160 GGACAGCCCTGACAGCCA 179

RESULT 7  
LOCUS AX026091 2380 bp DNA linear PAT 16-SEP-2000  
DEFINITION Sequence 3 from Patent DE19847779.  
ACCESSION AX026091  
VERSION AX026091.1 GI:10187522

KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Mueller-Schilling, M., Kramer, P. and Oren, M.  
TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy  
JOURNAL Patent: DE 19847779-C 3 03-FEB-2000;

DEUTSCHES KREBSFORSCH (DE)  
Location/Qualifiers  
1. .2380  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 579 a 595 c 568 g 638 t

ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 20; DB 6; Length 2380;  
100.0%; Pred. No. 26;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  
1 GGACAGCCCTGACAGCCA 20  
|||||  
1820 GGACAGCCCTGACAGCCA 1839

RESULT 8  
LOCUS AX026092 2827 bp DNA linear PAT 16-SEP-2000  
DEFINITION Sequence 4 from Patent DE19847779.  
ACCESSION AX026092  
VERSION AX026092.1 GI:10187523

KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Mueller-Schilling, M., Kramer, P. and Oren, M.  
TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy  
JOURNAL Patent: DE 19847779-C 4 03-FEB-2000;  
DEUTSCHES KREBSFORSCH (DE)

FEATURES  
source location/Qualifiers  
1. .2827  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 728 a 676 c 657 g 766 t

ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 20; DB 6; Length 2827;  
100.0%; Pred. No. 25;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  
1 GGACAGCCCTGACAGCCA 20  
|||||  
2267 GGACAGCCCTGACAGCCA 2286

RESULT 9  
LOCUS AX026089 3212 bp DNA linear PAT 16-SEP-2000  
DEFINITION Sequence 1 from Patent DE19847779.  
ACCESSION AX026089  
VERSION AX026089.1 GI:10187520

KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Mueller-Schilling, M., Kramer, P. and Oren, M.  
TITLE Novel receptor dna useful for identifying apoptosis-modulating substances potentially useful for cancer chemotherapy  
JOURNAL Patent: DE 19847779-C 1 03-FEB-2000;  
DEUTSCHES KREBSFORSCH (DE)

FEATURES  
source location/Qualifiers  
1. .3212  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"

```

BASE COUNT      778 a    784 c    809 g    841 t
ORIGIN
Query Match          100.0%; Score 20; DB 6; Length 3212;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGACAAAGCCCTGACAAGCCA 20
        |||||
Db       10653 GGACAAAGCCCTGACAAGCCA 10672

RESULT 11
LOCUS     AL157394
DEFINITION Human DNA sequence from clone RP11-399019 on chromosome 10, complete sequence.
ACCESSION AL157394
VERSION   AL157394.15 GI:15384622
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
AUTHORS   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE     1 (bases 1 to 187313)
JOURNAL   Blakey,S.
COMMENT   Direct Submissions
Submitted (22-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
On Aug 31, 2001 this sequence version replaced gi:1461146. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate

```

```

chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: EM, EMBL; SW,
SWISSPROT; Tr, TREMBL; Wp, WORMPEP; information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C.elegans/wormep
This sequence was generated from part of bacterial clone contigs of human
Chromosome 10, constructed by the Sanger Centre Chromosome 10
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chp10
RP11-399019 is from the library RCF1-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: PBACe3.6
This sequence is the entire insert of clone RP11-399019 The true
left end of clone RP11-496423 is at 166408 in this sequence. The
true right end of clone RP11-30415 is at 18704 in this sequence.

FEATURES
source
    location/Qualifiers
        1..187313
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="10"
            /clone="RP11-399019"
            /clone_1lb="RCF1-11.2"
            100119..100156
                /note="Sequence confirmed by AC015461 sequenced by W1BR."
            100157..100198
                /note="Sequence confirmed by AC015461 sequenced by W1BR."
            105808..105972
                /note="Sequence from AC015461 sequenced by W1BR."
            105973..105989
                /note="Sequence confirmed by AC015461 sequenced by W1BR."
BASE COUNT      55669 a 36398 c 36888 g 58358 t
ORIGIN
Query Match      100.0%; Score NO.; DB 9; Length 187313;
Best Local Similarity 100.0%; Pred. NO. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGACAAAGCCCTGACAAAGCCA 20
        |||||
        |||||
Db      144628 GGACAAAGCCCTGACAAAGCCA 144647

RESULT 12
AX026094 LOCUS      AX026094      20 bp      DNA      linear      PAT 16-SEP-2000
DEFINITION      Sequence 6 from Patent DE19847779.
ACCESSION      AX026094
VERSION      AX026094.1 GI:10187525
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Mueller-Schilling, M., Krammer, P. and Oren, M.
TITLE      Novel receptor dna useful for identifying apoptosis-modulating
            substances potentially useful for cancer chemotherapy
JOURNAL      Patent: DE 19847779-C 6 03-FEB-2000;
            DEUTSCHES KREBSFORSCH (DE)
FEATURES
source
    location/Qualifiers
        1..20
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
BASE COUNT      8 a 6 c 5 g 1 t
ORIGIN

```



Query Match 92.0%; Score 18.4; DB 6; Length 20;  
Best Local Similarity 95.0%; Pred. No. 4.5e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGACAAAGCCCTGACAAAGCA 20  
1 GGAAAAAGCCCTGACAAAGCA 20

RESULT 13  
ALJ36352 167412 bp DNA linear HTG 21-OCT-2001  
LOCUS Homo sapiens chromosome 1 clone RPS-859H16, \*\* SEQUENCING IN  
DEFINITION PROGRESS \*\*\*, 20 unordered pieces.  
ALJ36352  
ALJ36352.8 GI:16304915  
ACCESSION HTG; HTGS\_PHASE1; HTGS\_CANCELLED.  
VERSION HTG; HTGS\_PHASE1; HTGS\_CANCELLED.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
McLay, K.  
Direct Submission  
Submitted (18-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
On Oct 21, 2001 this sequence version replaced gi:12539659.  
----- Genome Center  
Center: Sanger Centre  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquerry@sanger.ac.uk  
Project Information  
Center project name: dJ859H16  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Sequencing vector: plasmid; 108752; 100% of reads  
Chemistry: Dye-terminator ET-amersham; 26% of reads  
Dye-terminator Big Dye; 73% of reads  
Consensus quality: 15949 bases at least Q40  
Consensus quality: 162148 bases at least Q30  
Consensus quality: 163926 bases at least Q20  
Insert size: 165512; sum-of-contigs  
Insert size: 206683; 10.2% error; agarose-fp  
Quality coverage: 4.08x in Q20 bases; sum-of-contigs Quality  
coverage: 3.39x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 20 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 12019: contig of 12019 bp in length  
\* 12020 12119: gap of 100 bp  
\* 12120 32657: contig of 20538 bp in length  
\* 32658 32757: gap of 100 bp  
\* 32758 36105: contig of 3348 bp in length  
\* 36106 36205: gap of 100 bp  
\* 36206 41069: contig of 4864 bp in length  
\* 41070 41169: gap of 100 bp  
\* 41170 51654: contig of 10485 bp in length  
\* 51655 51754: gap of 100 bp  
\* 51755 62397: contig of 10643 bp in length  
\* 62398 73595: gap of 100 bp  
\* 73596 73695: contig of 11098 bp in length  
\* 73696 83277: gap of 100 bp  
\* 83278 83377: contig of 9582 bp in length  
\* 83378 83377: gap of 100 bp

FEATURES  
source  
\* 83378 98845: contig of 15468 bp in length  
\* 98846 98945: gap of 100 bp  
\* 98946 101173: contig of 2228 bp in length  
\* 101174 101273: gap of 100 bp  
\* 101274 103420: contig of 2147 bp in length  
\* 103421 103520: gap of 100 bp  
\* 103521 109720: contig of 6200 bp in length  
\* 109721 109820: gap of 100 bp  
\* 109821 114256: contig of 4436 bp in length  
\* 114257 114356: gap of 100 bp  
\* 114357 119627: contig of 5271 bp in length  
\* 119628 119727: gap of 100 bp  
\* 119728 128350: contig of 8623 bp in length  
\* 128351 128450: gap of 100 bp  
\* 128451 140183: contig of 11733 bp in length  
\* 140184 140283: gap of 100 bp  
\* 140284 147703: contig of 7420 bp in length  
\* 147704 147803: gap of 100 bp  
\* 147804 158100: contig of 10297 bp in length  
\* 158101 158200: gap of 100 bp  
\* 158201 161982: contig of 3782 bp in length  
\* 161983 162082: gap of 100 bp  
\* 162083 167412: contig of 5330 bp in length.  
\* Location/Qualifiers  
1. 167412  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/clone="RPS-859H16"  
/clone\_11b="RPCT-5"  
1. 12019  
/note="assembly\_fragment:00380  
clone end:SP6  
vector\_side:left"  
12120. 32657  
/note="assembly\_fragment:00114  
fragment\_chain:1"  
32758. 36105  
/note="assembly\_fragment:00063  
fragment\_chain:1"  
36206. 41069  
/note="assembly\_fragment:01365  
fragment\_chain:1"  
41170. 51654  
/note="assembly\_fragment:01782  
fragment\_chain:1"  
51735. 62397  
/note="assembly\_fragment:01003  
fragment\_chain:1"  
62498. 73595  
/note="assembly\_fragment:00248  
fragment\_chain:2"  
73696. 83277  
/note="assembly\_fragment:01125  
fragment\_chain:2"  
83378. 98845  
/note="assembly\_fragment:01780  
fragment\_chain:2"  
98946. 101173  
/note="assembly\_fragment:01201  
fragment\_chain:2"  
101274. 103420  
/note="assembly\_fragment:01704  
fragment\_chain:3"  
103521. 109720  
/note="assembly\_fragment:00350  
fragment\_chain:3"  
109821. 114256  
/note="assembly\_fragment:01435  
fragment\_chain:3"  
114357. 119627  
/note="assembly\_fragment:00291

```

misc_feature      119728.0 128350.0 /note="assembly_fragment:01405  

fragment_chain:4"  

misc_feature      128451.0 140183.0 /note="assembly_fragment:00901  

fragment_chain:5"  

misc_feature      140284.0 147703.0 /note="assembly_fragment:00215  

fragment_chain:5"  

misc_feature      147804.0 158100.0 /note="assembly_fragment:00322"  

fragment_chain:5"  

misc_feature      158201.0 161982.0 /note="assembly_fragment:00595"  

fragment_chain:5"  

misc_feature      162083.0 167412.0 /note="assembly_fragment:00872"  

fragment_chain:5"  

misc_feature      1904.0 1904.0 /note="assembly_fragment:00872"  

fragment_chain:5"  


```

QY	1	GGACAGCCTGACAGCCA	20
Query Match			92.0%
Best Local Similarity			95.0%
Matches	19;	Conservative	0;
		Mismatches	1;
		Indels	0;
		Gaps	0;
Score	18.4;	DB	2;
Length	167412;		
Pred. No. 70;			
DB	143124	GGTCAGGCCCTGACAGCCA	143105

RESULT 14				
AL672160				
LOCUS	AL672160	192657 bp	DNA	linear
DEFINITION	Mus DNA sequence from clone RP23-144K18 on chromosome 4, complete mouse.			
ACCESSION	AL672160			
VERSION	AL672160.10	GI:25809448		
KEYWORDS	HTG.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
	1 (bases 1 to 192657)			
	Hammond, S.			
	Direct Submission			
REFERENCE	Submitted (17-APR-2003) Wellcome Trust Sanger Institute, Hinxton,			
AUTHORS	Cambriergeshtre, CB10 1SA, UK. E-mail enquiries:			
TITLE	humgery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk			
JOURNAL	On Nov 27, 2002 this sequence version replaced gi:24394955.			
COMMENT				

a printed quality of at least 30.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)  
 -----

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

```

FEATURES
source
    1.192657
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    /chromosome="4"
    /clone="RP23-144K18"
    /clone_1b="RPCT-23"
    location/Qualifiers
BASE COUNT
51940 a 39496 c 42716 g 58505 t
ORIGIN

```

	92.0%;	Score 18.4;	DB 10;	Length 192657;
Query Match	Best Local Similarity	95.0%;	Pred. No. 68;	
Matches 19; Conservative	0;	Mismatches	1;	Indels 0;
Gaps	0;			
QY	1	GGACAGCCCTGACAAGCA	20	
.				
Db	156517	GGACAGCCCGACACAAGCA	156536	

RESULT 15				
BX294664/c				
LOCUS	BX294664	193152 bp	DNA	linear HTG 24-APR-2003
DEFINITION	Mus musculus chromosome X clone RP23-384C22, *** SEQUENCING IN PROGRESS ***, 10 unordered pieces.			
ACCESSION	BX294664			
VERSION	BX294664.2	GI:29500936		
KEYWORDS	HTG; HTGS PHASEI; HTGS CANCELLED.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
	1 (bases 1 to 193152)			
	McLay,K.			
REFERENCE	Direct Submission			
AUTHORS	Submitted (23-APR-2003) Wellcome Trust Sanger Institute, Hinxton,			
TITLE	Cambridgeshire, Chr10 ISA, UK. E-mail enquiries:			
JOURNAL	humgen@wanger.ac.uk Clone requests: clonerequest@sanger.ac.uk			
COMMENT	On Apr 2, 2003 this sequence version replaced gi:29134752.			

```

-----
Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humgureys@sanger.ac.uk
-----
Project Information
Center project name: DM384C22
-----
Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 190946 bases at least Q40
Consensus quality: 191521 bases at least Q30
Consensus quality: 191867 bases at least Q20
Insert size: 192452; sum-of-contigs
Insert size: 197663; 4.4% error; agarose-fp
Quality coverage: 8.26x in Q20 bases; sum-of-contigs Quality
coverage: 8.09x in Q20 bases; agarose-fp
-----

```

\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 37461: contig of 37461 bp in length  
\* 37462 37561: gap of 100 bp  
\* 37562 45950: contig of 8389 bp in length  
\* 45951 46051: gap of 100 bp  
\* 46051 98426: contig of 52376 bp in length  
\* 98427 98526: gap of 100 bp  
\* 98527 118027: contig of 19501 bp in length  
\* 118028 118127: gap of 100 bp  
\* 118128 126520: contig of 8393 bp in length  
\* 126521 132811: gap of 100 bp  
\* 132812 132911: contig of 6191 bp in length  
\* 132912 172432: gap of 100 bp  
\* 172433 172532: contig of 39521 bp in length  
\* 172533 185735: gap of 100 bp  
\* 185736 185835: contig of 1303 bp in length  
\* 185836 190693: gap of 100 bp  
\* 190694 190793: contig of 4858 bp in length  
\* 190794 193352: gap of 100 bp  
\* 193353 193352: contig of 2559 bp in length.  
Location/Qualifiers  
1. 193352

FEATURES  
source

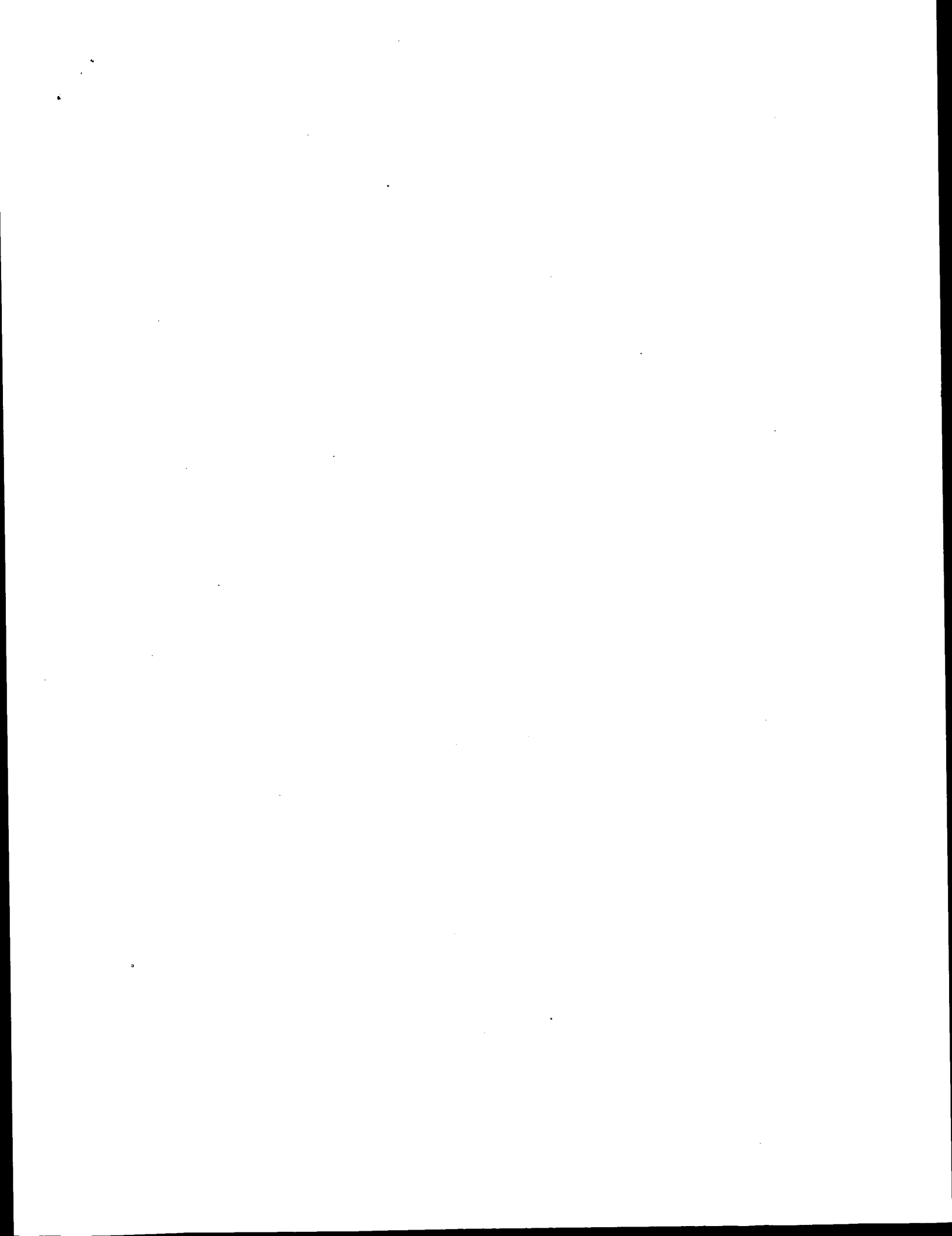
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="X"  
/clone="RP23-384C22"  
1. 37461  
/clone\_id="RPCI-23"  
/note="assembly\_fragment:02244  
fragment\_chain:1  
clone\_end:SP6  
vector\_side:left"  
37562. 45950  
/note="assembly\_fragment:00826  
fragment\_chain:1"  
46051. 58426  
/note="assembly\_fragment:00297.  
fragment\_chain:1"  
98527. 118027  
/note="assembly\_fragment:01245  
fragment\_chain:1"  
118128. 126520  
/note="assembly\_fragment:02356  
fragment\_chain:1"  
126621. 132811  
/note="assembly\_fragment:01763  
fragment\_chain:1"  
132912. 172432  
/note="assembly\_fragment:01130  
fragment\_chain:1"  
172533. 185735  
/note="assembly\_fragment:01150  
fragment\_chain:1"  
185836. 190693  
/note="assembly\_fragment:00830  
fragment\_chain:1  
clone\_end:T7  
vector\_side:right"  
190794. 193352  
/note="assembly\_fragment:00599"  
42616 a 42616 c 39284 g 52100 t 900 others

BASE COUNT 58452 a 42616 c 39284 g 52100 t 900 others  
ORIGIN

Query Match 92.0%; Score 18.4; DB 2; Length 193352;  
Best Local Similarity 95.0%; Pred. No. 68;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACAGGCTGTGACAGCCA 20  
DB 36135 GGACAGGCCGACAGCCA 36116

Search completed: August 1, 2003, 17:33:04  
Job time : 80.1573 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:13:22 ; Search time 68.4754 Seconds  
(without alignments)  
7098.748 Million cell updates/sec

Title: US-09-834-291-24

Sequence: 1 ggacaagcccgacaagcca 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estda:\*  
2: em\_estin:\*  
3: em\_estin:\*  
4: em\_estin:\*  
5: em\_estin:\*  
6: em\_estin:\*  
7: em\_estin:\*  
8: em\_estin:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estfun:\*  
17: em\_estfun:\*  
18: em\_estfun:\*  
19: em\_estfun:\*  
20: em\_estfun:\*  
21: em\_estfun:\*  
22: em\_estfun:\*  
23: em\_estfun:\*  
24: em\_estfun:\*  
25: em\_estfun:\*  
26: em\_estfun:\*  
27: em\_estfun:\*  
28: gb\_est1:\*  
29: gb\_est2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	793	12	BI763679
2	20	100.0	962	9	AL540709
3	20	100.0	1089	13	BO072834
4	18	90.0	486	14	W45225

Result No.	Score	Query Match	Length	DB ID	Description
5	18	90.0	506	10	BG659530
6	18	90.0	753	13	BUI42933
7	17.4	87.0	619	28	A2939002
8	17.4	87.0	684	28	A2450876
9	17.4	87.0	689	29	BX220214
10	17.4	87.0	701	14	CA605171
11	17.4	87.0	745	29	BX179961
12	17	85.0	331	14	TE2311
13	17	85.0	429	28	A2781046
14	16.8	84.0	531	9	AI195416
15	16.8	84.0	592	9	AW131236
16	16.8	84.0	629	28	A2492025
17	16.8	84.0	730	28	BH988275
18	16.8	84.0	766	13	B0471804
19	16.8	84.0	843	14	CB560391
20	16.8	84.0	973	13	B0715536
21	16.8	84.0	1059	10	BF205338
22	16.8	84.0	1077	14	CD519542
23	16.8	84.0	1058	11	AK043145
24	16.8	82.0	191	10	BF386638
25	16.4	82.0	262	28	AC632653
26	16.4	82.0	312	12	BI041486
27	16.4	82.0	424	13	BY444129
28	16.4	82.0	545	10	BG376805
29	16.4	82.0	626	28	BH303980
30	16.4	82.0	652	29	AG094812
31	16.4	82.0	675	28	AG066797
32	16.4	82.0	696	10	BF966973
33	16.4	82.0	742	29	BZ322825
34	16.4	82.0	802	10	BG481356
35	16.4	82.0	806	10	BG296720
36	16.4	82.0	809	28	AQ050636
37	16.4	82.0	890	14	CB192716
38	16.4	82.0	926	13	BH849141
39	16.4	82.0	1114	29	CC278162
40	16.4	82.0	2193	10	BF579152
41	16	80.0	300	9	AL916663
42	16	80.0	406	14	CB807904
43	16	80.0	407	14	CB771883
44	16	80.0	417	14	CB766812
45	16	80.0	427	14	CB795385

## ALIGNMENTS

RESULT 1  
BI763679/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BI763679 793 bp mRNA linear EST 25-SEP-2001  
603049567F1 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5189752.5,  
mRNA sequence.  
BI763679  
BI763679.1 GI:15755257  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
NIH-MGC http://mgs.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L14M11474 row: e column: 17  
High quality sequence stop: 786.

# FEATURES

## source

Location/Qualifiers  
1. .793  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5189752"  
/lab\_host="DH10B"  
/clone\_1lb="NIH MGC 116"  
/note="Organ: pooled colon, kidney, stomach; Vector: PCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH MGC library."

BASE COUNT 183 a 265 c 208 g 137 t

ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 793;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20  
|||||  
Db 398 GGACAGCCCTGACAGCCA 379

RESULT 2  
AL540709/c 962 bp mRNA linear EST 12-MAY-2003  
LOCUS AL540709 Homo sapiens PLACENTA Homo sapiens CDNA clone CS0DE002YN18  
DEFINITION 5-PRIME, mRNA sequence.  
ACCESSION AL540709  
VERSION AL540709.2 GI:30544172  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 962)  
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
On Feb 15, 2001 this sequence version replaced gi:12871113.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr; Web: www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 5554.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DE002DG09P1&cluster=5554.r. Contact :  
Peng Liang Email : lliang@lifestech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DE002DG09P1.

# FEATURES

## source

Location/Qualifiers  
1. .962  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DE002YN18"  
/tissue\_type="PLACENTA"  
/note="Vector: PCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the PCMVSPORT 6 vector.  
Library was not normalized."

BASE COUNT 211 a 302 c 264 g 181 t 4 others

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 962;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20  
|||||  
Db 94 GGACAGCCCTGACAGCCA 75

RESULT 3  
B0072834 1089 bp mRNA linear EST 02-APR-2002  
LOCUS AGENCOURT 6763280 NIH\_MGC\_118 Homo sapiens CDNA clone IMAGE:5756324  
DEFINITION 5', mRNA sequence.  
ACCESSION B0072834 GI:19901880  
VERSION B0072834.1 GI:19901880  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1089)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Agencourt Bioscience Corporation  
Clone distributing: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM2796 row: 1 column: 21  
High quality sequence stop: 645.

# FEATURES

## source

Location/Qualifiers  
1. .1089  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5756324"  
/tissue\_type="leukocyte"  
/lab\_host="DH10B"  
/clone\_1lb="NIH MGC 118"  
/note="Vector: PCMV-SPORT6; Site 1: NotI; Site 2: EcoRV  
(destroyed); RNA source leukocytes from anonymous pool of  
non-activated adult donors. Library is oligo-dT primed  
and directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.7 kb, insert size range  
1.2-3.3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 027. Note:  
this is a NIH MGC library."

BASE COUNT 280 a 263 c 333 g 212 t 1 others

ORIGIN

Query Match 100.0%; Score 20; DB 13; Length 1089;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20  
|||||  
Db 369 GGACAGCCCTGACAGCCA 388

RESULT 4  
W45225 486 bp mRNA linear EST 10-OCT-1996  
LOCUS W45225  
DEFINITION zc33c09.r1 Soares\_senescent\_fibroblasts\_NbHSF Homo sapiens CDNA

clone IMAGE:323152.5, similar to gb:125081 TRANSFORMING PROTEIN  
 RHOC (HUMAN); mRNA sequence.  
 ACCESSION W45225  
 VERSION W45225.1 GI:1329306  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 486)  
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman  
 M., Hiltman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,  
 Rife, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston  
 R., Williamson, A., Woldmann, P., and Wilson, R.  
 The WashU-Merck EST Project  
 JOURNAL Unpublished  
 COMMENT Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.wustl.edu  
 This clone is available royalty-free through INM; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert length: 972 Std Error: 0.00  
 Seq primer: mob.RBGA+BT  
 High quality sequence stop: 296.  
 Location/Qualifiers  
 1. 486  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:1254664"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:323152"  
 /tissue\_type="senescent fibroblast"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares senescent fibroblasts NDHSF"  
 /note="Vector: pT73D (Pharmacia) with a modified  
 polylinker V-type phagemid; Site 1: Not I; Site 2: Eco RI  
 ; 1st strand cDNA was primed with a Not I - oligo(dT)  
 primer [5',  
 TGTTACCAATCTGAACTGGAGCGCCGCAATTTTCTTTTCTTTT 3']  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pT73 vector  
 (Pharmacia). Library went through one round of  
 normalization to a Cot = 5. Library constructed by Bento  
 Soares and M. Fatima Bonaldo."  
 BASE COUNT 120 a 129 c 140 g 95 t 2 others  
 ORIGIN  
 Query Match 90.0%; Score 18; DB 14; Length 486;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 GACAGCCCTGACAGCC 19  
 DB 265 GACAGCCCTGACAGCC 282  
 RESULT 5  
 BG659530 506 bp mRNA linear EST 11-MAY-2001  
 LOCUS BG659530  
 DEFINITION TIGEST24901.Y1 TVEG118 Tachyzoite cDNA library Toxoplasma gondii  
 CNA clone TIGEST24901.Y1 5' similar to TR:063030 063030 RAT  
 ALPNA-SMOOTH MUSCLE ACTIN mRNA FRAGMENT; mRNA sequence.  
 ACCESSION BG659530  
 VERSION BG659530.1 GI:13801991  
 KEYWORDS EST.  
 SOURCE Toxoplasma gondii  
 ORGANISM Toxoplasma gondii  
 Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;  
 Sarcocystidae; Toxoplasma.

REFERENCE 1 (bases 1 to 506)  
 Tang, K., Cole, R., Fogarty, S., Sibley, L.D., Ajika, J.A., White, M.,  
 Clifton, S., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M.,  
 Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Riter  
 E., Bennett, J., Franklin, C., Tsagaris, R., Ronko, I., Kennedy  
 S., Maguire, L., Waterston, R., and Wilson, R.  
 Toxoplasma EST Project  
 JOURNAL Unpublished  
 COMMENT Contact: Clifton, S.  
 Toxoplasma EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: tox@wustl.wustl.edu  
 Contact David Sibley (tox@est@wustl.wustl.edu) for further  
 information relating to organism, libraries, or clone availability.  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 371.  
 Location/Qualifiers  
 1. 506  
 /organism="Toxoplasma gondii"  
 /mol\_type="mRNA"  
 /strain="VEG"  
 /db\_xref="taxon:5811"  
 /clone="TIGEST24901.Y1"  
 /dev\_stage="Tachyzoite"  
 /lab\_host="DH10B"  
 /clone\_lib="TVEG118 Tachyzoite cDNA library"  
 /note="Vector: pRiuescript SK; Site 1: EcoRI; Site 2: XhoI  
 ; This library was constructed by Kellian Tang, Robert  
 Cole, and L. David Sibley at Washington University. cDNAs  
 were synthesized from poly(A) + RNA by oligo(dT) priming,  
 size-selected and directionally cloned into the uni-zap XR  
 lambda vector (Stratagene). The primary library was mass  
 excised as phagemids and rescued in SOLR cells. The  
 plasmid library was recovered from the SOLR cells and  
 transformed in mass into DH10B (GeneHog, Research Genetics  
 , Inc.) for sequencing. WARNING: This library may contain  
 a small percentage contaminants from human fibroblast  
 cells."

BASE COUNT 86 a 191 c 140 g 89 t  
 ORIGIN  
 Query Match 90.0%; Score 18; DB 10; Length 506;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGACAGCCCTGACAGC 18  
 DB 18 GGACAGCCCTGACAGC 1  
 RESULT 6  
 BU142933 753 bp mRNA linear EST 25-NOV-2002  
 LOCUS BU142933  
 DEFINITION 603135537F1 CSRGCHL25 Gallus gallus cDNA clone CHEST1919.5', mRNA  
 sequence.  
 ACCESSION BU142933  
 VERSION BU142933.1 GI:25358894  
 KEYWORDS EST.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 REFERENCE 1 (bases 1 to 753)  
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,  
 Ford, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A., and Hubbard, S.J.  
 A Comprehensive Collection of Chicken cDNAs  
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)  
 MEDLINE 2235534  
 PUBMED 12445392





## FEATURES

source

Seq primer: CACACAGGAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 684.  
 Location/Qualifiers

1..684  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUCG1M0249M17"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid UUCG1M library"  
 /note="Vector: PMD42uv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## BASE COUNT

180 a 202 c 160 g 142 t

## ORIGIN

Query Match 87.0%; Score 17.4; DB 28; Length 684;  
 Best Local Similarity 94.7%; Pred. No. 7.3e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACAGCCCTGACAGCCA 20

Db 611 GACAGCAGCTGACAGCCA 629

RESULT 9  
 LOCUS BX220214/c 689 bp DNA linear GSS 29-JAN-2003  
 DEFINITION Danio rerio genomic clone DKEX-269M11, genomic survey sequence.  
 ACCESSION BX220214  
 VERSION BX220214.1 GI:28052100  
 KEYWORDS GSS:  
 SOURCE Danio rerio (zebrafish)  
 ORGANISM Danio rerio  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.  
 1 (bases 1 to 689)  
 Humphrey,S.J., Huckle,E. and Durham,J.L.  
 Direct Submission  
 Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humphrey@sanger.ac.uk Unpublished  
 This sequence was generated from the SP6 end of BAC 269M11. 269M11 is part of the Daniokey BAC library created by R. Plasterk and N.V. Keygene. Further details:  
 http://www.sanger.ac.uk/Projects/D\_rerio/  
 Location/Qualifiers

## FEATURES

source

1..689  
 /organism="Danio rerio"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /clone="DKEX-269M11"  
 /tissue\_type="Testis"

BASE COUNT 180 a 144 c 144 g 221 t  
 /note="vector pIndigoBAC-536"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 29; Length 689;  
 Best Local Similarity 94.7%; Pred. No. 7.4e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GACAGCCCTGACAGCCA 20

Db 419 GACAGCCCTGACAGCCA 401

RESULT 10  
 LOCUS CA605171/c 701 bp mRNA linear EST 21-NOV-2002  
 DEFINITION wrl.pk0046.e7 wrl Triticum aestivum cDNA clone wrl.pk0046.e7 5' end  
 mRNA sequence.  
 ACCESSION CA605171  
 VERSION CA605171.1 GI:25160333  
 KEYWORDS EST.  
 SOURCE Triticum aestivum (bread wheat)  
 ORGANISM Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae  
 / Triticaceae; Triticum.  
 1 (bases 1 to 701)  
 Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z., Miao,G., Caraher,N. and Hanafey,M.K.  
 Dupont Wheat cDNA Sequence  
 Unpublished  
 Contact: Scott V. Tingey  
 Crop Genetics  
 E. I. Dupont de Nemours and Company  
 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA  
 Tel: 302-631-2602  
 Fax: 302-631-2607  
 Email: Scott.V.Tingey@usa.dupont.com  
 Seq primer: M13.  
 Location/Qualifiers

## FEATURES

source

1..701  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4565"  
 /clone="wrl.pk0046.e7"  
 /tissue\_type="root"  
 /clone\_lib="wrl"  
 /note="Vector: pBluescript SK+; Site\_1: EcoRI; Site\_2: XhoI; Wheat (Triticum aestivum L.) root; 7 day old seedling, light grown"  
 BASE COUNT 199 a 186 c 131 g 156 t 27 others  
 ORIGIN

Query Match 87.0%; Score 17.4; DB 14; Length 701;  
 Best Local Similarity 94.7%; Pred. No. 7.4e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCC 19

Db 255 GGACAGCCCTGACAGCC 237

RESULT 11  
 LOCUS BX179961/c 745 bp DNA linear GSS 28-JAN-2003  
 DEFINITION Danio rerio genomic clone DKEX-183M20, genomic survey sequence.  
 ACCESSION BX179961  
 VERSION BX179961.1 GI:28011764  
 KEYWORDS GSS:  
 SOURCE Danio rerio (zebrafish)  
 ORGANISM Danio rerio  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

REFERENCE	Cypriniformes; Cyprinidae; Danio.
AUTHORS	1 (bases 1 to 745)
TITLE	Humphray, S.J., Huckle, E. and Durham, J.L.
JOURNAL	Direct Submission Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgehire, CB10 1SA, UK. E-mail enquiries: humq@anger.ac.uk Unpublished This sequence was generated from the SP6 and of BAC 183M20. 183M20 is part of the Daniokey BAC library created by R. Plastek and N.V. Keygene. Further details: <a href="http://www.sanger.ac.uk/projects/D_reio/">http://www.sanger.ac.uk/projects/D_reio/</a> .
COMMENT	
FEATURES	Location/Qualifiers
source	1..745
	/organism="Danio rerio"
	/mol_type="genomic DNA"
	/db_xref="taxon:7955"
	/clone="DKEXY-183M20"
	/tissue_type="Testis"
	/note="vector pindigOBAC-536"
BASE COUNT	209 a 150 c 151 g 235 t
ORIGIN	
Query Match	87.0%; Score 17.4; DB 29; Length 745;
Best Local Similarity	94.7%; Pred. No. 7.6e+02;
Matches	18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	2 GACAAAGCCCTGACAAGCCA 20
Db	419 GACAAAGCCCTGACAAGCCA 401
RESULT 12	
LOCUS	T62311
DEFINITION	T62311
ACCESSION	EGEST0074 TGRH Tachyzoite cDNA Toxoplasma gondii cDNA clone tgb008
VERSION	5', end, mRNA sequence.
KEYWORDS	T62311
SOURCE	T62311.1 GI:665753
ORGANISM	EST.
	Toxoplasma gondii
	Toxoplasma gondii
	Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
	Sarcocystidae; Toxoplasma.
	1 (bases 1 to 331)
	Man, K.-L., Blackwell, J.M. and Ajioke, J.W.
	Toxoplasma gondii expressed sequence tags: insight into tachyzoite
	gene expression
	Mol. Biochem. Parasitol. 75 (2), 179-186 (1995)
JOURNAL	95258558
MEDLINE	8992316
PUBMED	
COMMENT	Contact: Ajioke JM
	Laboratory for Parasite Genome Analysis
	Cambridge University
	Department of Pathology, Tennis Court Road, Cambridge CB2 1QP, UK.
	Tel: 01223333923
	Fax: 01223333923
	Email: jajioke@gmp.mrc.ac.uk
	Clone information and library can be obtained from Jim Ajioke
	jwasmole.bio.cam.ac.uk
	Seq primers: SK;
FEATURES	Location/Qualifiers
source	1..331
	/organism="Toxoplasma gondii"
	/mol_type="mRNA"
	/strain="RH"
	/db_xref="taxon:5811"
	/clone="tgb008"
	/lab_host="XLI-Blue MRF"
	/clone_lib="TGRH Tachyzoite cDNA"
	/note="Vector: lambda ZAP; Site 1: EcoRI; Site 2: XhoI;
	Toxoplasma RH strain tachyzoites were grown in human
	foreskin fibroblast cultures in vitro. The library was
	constructed by K.L. Man, Cambridge University. cDNAs were

synthesized from polyA RNAs by oligo d(T) priming and directionally cloned into the EcoRI to XhoI sites of the Lambda ZapII vector using the Zap-CDNA synthesis kit (Stratagene). WARNING: the library contains a small percentage of cDNAs derived from the human host cells."

BASE COUNT	70 a	94 c	80 g	87 t
ORIGIN				
Query Match	85.0%;	Score 17;	DB 14;	Length 331;
Best Local Similarity	100.0%;	Pred. No. 8,88+02;		
Matches	17;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0
QY	1 GGACACAGCCCTGACAAAG 17			
Db	314 GGACACAGCCCTGACAAAG 330			
RESULT 13				
AZ781046				
LOCUS				
DEFINITION	AZ781046 429 bp DNA linear GSS 16-FEB-2001			
ACCESSION	2M0018K23R Mouse 10kb plasmid UUGC1M library Mus musculus genomic			
VERSION	clone UUGC2M0018K23 R, genomic survey sequence.			
KEYWORDS	AZ781046.1 GI:12913345			
SOURCE	GSS.			
ORGANISM	Mus musculus (house mouse)			
REFERENCE	Mus musculus			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 429)			
JOURNAL	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D.,Weiss,R.			
COMMENT	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts			
UNPUBLISHED				
CONTACT:	Robert B. Weiss			
UNIVERSITY OF UTAH	Genome Center			
RM. 308, BIOMEDICAL POLYMERS RESEARCH BLDG., 20 S. 2030 E., SLC, UT 84112, USA				
TEL: 801 585 5606				
FAX: 801 585 7177				
EMAIL: ddunn@genetics.utah.edu				
INSERT LENGTH: 10000	Std Error: 0.00			
PLATE: 0018	row: K column: 23			
SEQ PRIMER: CACACAGAAACAGCATATGACC				
CLASS: plasmid ends				
High quality sequence stop: 429.				
LOCATION/Qualifiers				
1. 429				
/organism="Mus musculus"				
/mol_type="genomic DNA"				
/strain="C57BL/6J"				
/db_xref="taxon:10090"				
/clone="UUGC2M0018K23"				
/sex="Male"				
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"				
/clone_lib="Mouse 10kb plasmid UUGC1M library"				
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource				
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi 4732114 gb AF159072.1), a copy-number				

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 125 a 88 c 126 g 90 t

ORIGIN

Query Match 85.0%; Score 17; DB 28; Length 429;  
Best Local Similarity 100.0%; Pred. No. 9.6e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGGCGCTGACAG 17  
227 GGACAGGCGCTGACAG 243

DB

RESULT 14  
AA195416/c  
LOCUS AA195416 531 bp mRNA linear EST 06-AUG-1997  
DEFINITION x16f10.s1 Soares\_NHMPu\_S1 Homo sapiens cDNA clone IMAGE:665515  
3', mRNA sequence.  
ACCESSION AA195416  
VERSION AA195416.1 GI:1785109  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE

ATTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kucab, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R., and Wilson, R.  
WashU-Merck EST Project 1997  
Unpublished

TITLE  
JOURNAL  
COMMENT  
Contact: Wilson RK  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available royalty-free through LNU; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 1020 Std Error: 0.00  
High quality sequence stop: 436.

FEATURES

source

1..531  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:5427547"  
/db\_xref="taxon:9606"  
/clone="IMAGE:665515"  
/tissue\_type="Pooled human melanocyte, fetal heart, and pregnant uterus"  
/lab\_host="DH10B"  
/clone\_lib="Soares\_NHMPu\_S1"  
/note="Organ: mixed (see below); Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NHMP, pregnant uterus 2NHMP, and fetal heart 2NHMP) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT 111 a 110 c 167 g 141 t 2 others

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 531;

Best Local Similarity 90.0%; Pred. No. 1.3e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAGGCGCTGACAGCCA 20  
158 GGACAGGCGCTGACAGCCA 139

DB

RESULT 15  
AM131236/c  
LOCUS AM131236 592 bp mRNA linear EST 27-OCT-1999  
DEFINITION x160c04.x1 NCI CGAP Gas4 Homo sapiens cDNA clone IMAGE:2622438 3', similar to SW:TP1B HUMAN Q13263 TRANSCRIPTION INTERMEDIARY FACTOR 1-BETA', mRNA sequence.  
ACCESSION AM131236  
VERSION AM131236.1 GI:6132843  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE

ATTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kucab, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R., and Wilson, R.  
WashU-Merck EST Project 1997  
Unpublished

TITLE  
JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
cDNA distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNU at: www.bio.llnl.gov/bdrrp/image/image.html  
Seq primer: -400P from Gibco  
High quality sequence stop: 423.

FEATURES

source

1..592  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2622438"  
/tissue\_type="poorly differentiated adenocarcinoma with signer ring cell features"  
/lab\_host="DH10B"  
/clone\_lib="NCI-CGAP Gas4"  
/note="Organ: stomach; Vector: PCMV-SPORE6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"

BASE COUNT 133 a 146 c 194 g 115 t 4 others

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 592;  
Best Local Similarity 90.0%; Pred. No. 1.3e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAGGCGCTGACAGCCA 20  
590 GGACAGGCGCTGACAGCCA 571

DB

Search completed: August 1, 2003, 20:43:07  
Job time: 68.4754 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:32:17 ; Search time 6.6986 Seconds  
(without alignments)  
8059.612 Million cell updates/sec

Title: US-09-834-291-24

Perfect score: 20  
Sequence: 1 ggacaagccctgacaaagca 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_190un03.\*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	266	21	AAZ88700
2	16.4	82.0	516	25	ABX81757
3	16.4	82.0	8243	25	ABZ74553
4	16.4	82.0	12710	22	AAK81012
5	16.4	82.0	12710	25	ABZ74552
6	16	80.0	433	25	ABX52814
7	15.8	79.0	165	25	ABX27459
8	15.8	79.0	261	25	ABZ41531

C	9	15.8	79.0	273	20	AAZ11974
C	10	15.8	79.0	288	21	AAZ53178
C	11	15.8	79.0	309	21	AAZ53176
C	12	15.8	79.0	381	20	AAZ11975
C	13	15.8	79.0	409	22	AAK58992
C	14	15.8	79.0	749	22	AAK74242
C	15	15.8	79.0	749	22	AAK74243
C	16	15.8	79.0	750	22	AAK74241
C	17	15.8	79.0	1012	23	AAZ59679
C	18	15.8	79.0	1482	25	ACA01448
C	19	15.8	79.0	1482	25	ACA01793
C	20	15.8	79.0	16949	23	AAZ59585
C	21	15.8	79.0	39325	22	AAK81660
C	22	15.8	79.0	88421	24	AAZ40781
C	23	15.8	79.0	136328	24	ABZ35015
C	24	15.8	79.0	155074	24	AAH85735
C	25	15.8	79.0	309400	22	AAH68534
C	26	15.4	77.0	272	25	ABX54046
C	27	15.4	77.0	401	22	AAK95694
C	28	15.4	77.0	401	22	AAK97187
C	29	15.4	77.0	401	24	ABT00464
C	30	15.4	77.0	401	24	ABT01957
C	31	15.4	77.0	498	24	ABN68645
C	32	15.4	77.0	673	22	AAZ35624
C	33	15.4	77.0	673	25	ABX58612
C	34	15.4	77.0	758	20	AAZ16661
C	35	15.4	77.0	845	24	AAZ62770
C	36	15.4	77.0	860	22	AAZ37560
C	37	15.4	77.0	860	22	AAZ37561
C	38	15.4	77.0	860	25	ABX60549
C	39	15.4	77.0	860	25	ABX60549
C	40	15.4	77.0	921	21	AAZ93516
C	41	15.4	77.0	1560	21	AAZ9701
C	42	15.4	77.0	2010	21	AAZ96604
C	43	15.4	77.0	2059	21	AAZ59970
C	44	15.4	77.0	3038	25	ABT17900
C	45	15.4	77.0	3604	25	ABT19714

## ALIGNMENTS

RESULT 1  
AAZ88700 standard; DNA; 266 BP.

AAZ88700	AAZ88700	standard; DNA; 266 BP.
XX	XX	
AC	AAZ88700;	
XX	XX	
DT	11-MAY-2000	(first entry)
XX	XX	
DE	Human CD95 receptor intron 1 fragment.	
XX	XX	
KM	p53; CD95 receptor; human; screening; apoptosis-modulation;	
KW	Cancer chemotherapy; ss.	
XX	XX	
OS	Homo sapiens.	
XX	XX	
FT	Key	Location/Qualifiers
FT	protein_bind	160..179
FT		/*tag= a
FT		/bound_molety= p53
XX	XX	
PN	DE19847779-Cl.	
XX	XX	
PD	03-FEB-2000.	
XX	XX	
PF	16-OCT-1998;	98DE-1047779.
XX	XX	
PR	16-OCT-1998;	98DE-1047779.
XX	XX	
PA	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.	
XX	XX	
PI	Krammer P, Mueller-Schilling M, Oren M;	

Neisseria meningit  
Neisseria meningit  
Neisseria gonorrhoe  
Neisseria gonorrhoe  
Human immune/haema  
Human immune/haema  
Human immune/haema  
Human immune/haema  
Propionibacterium  
C. glutamicum deri  
C. glutamicum deri  
Propionibacterium  
Human immune/haema  
8842Int genomic DN  
Human gene express  
Human genomic regi  
C glutamicum codin  
Bovine EST associa  
Human neutregulin g  
Human neutregulin g  
Human neutregulin-1  
Human neutregulin-1  
Streptococcus poly  
Human musculoskele  
CDNA encoding nove  
CDNA encoding nove  
Human gene express  
CDNA sequence #557  
Human musculoskele  
Human musculoskele  
CDNA encoding nove  
CDNA encoding nove  
Human secreted pro  
Human secreted pro  
Human secreted pro  
Aspergillus fumiga  
Aspergillus fumiga

XX DR WPI; 2000-162245/15.  
 XX PT Novel receptor DNA useful for identifying apoptosis-modulating  
 PT substances potentially useful for cancer chemotherapy  
 XX Claim 2; Fig 4; 12pp; German.  
 XX CC This invention describes a novel p53-binding region of a human CD95  
 CC receptor DNA molecule. The p53-binding region, or a vector containing  
 CC it, can be used to screen for apoptosis-modulating substances  
 CC potentially useful for cancer chemotherapy. This sequence represents a  
 CC fragment of the human CD95 receptor intron 1 which contains a p53 binding  
 CC region described in the method of the invention.  
 XX SQ Sequence 266 BP; 49 A; 72 C; 110 G; 35 T; 0 other;  
 Query Match 100.0%; Score 20; DB 21; Length 266;  
 Best Local Similarity 100.0%; Pred. No. 1.9;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGACAAAGCCCTGACAGCCA 20  
 Db 160 GGACAAAGCCCTGACAGCCA 179  
 RESULT 2  
 ABX91757/c  
 ID ABX91757 standard; cDNA; 516 BP.  
 XX AC ABX91757;  
 XX DT 07-MAY-2003 (first entry)  
 XX DE Murine gene trapped sequence (GTS) SEQ ID NO 1101.  
 XX DE Murine; mouse; gene trap technology; gene trapped sequence; GTS;  
 KW gene identification; functional genomic analysis; gene discovery;  
 KW gene expression analysis; cross species hybridisation analysis;  
 KW antisense inhibition; gene targeting; gene; ss.  
 XX OS Mus sp.  
 XX PN US2002161207-A1.  
 XX PD 31-OCT-2002.  
 XX PF 30-NOV-2000; 2000US-0728444.  
 XX PR 01-DEC-1999; 99US-168360P.  
 XX PA (FRIE/) FRIEDRICH G.  
 XX PA (ZAMB/) ZAMBROWICZ B.  
 XX PA (SAND/) SANDS A T.  
 XX PI Friedrich G, Zambrowicz B, Sands AT;  
 XX WPI; 2003-288124/28.  
 XX DR WPI; 2003-288124/28.  
 XX PT New murine polynucleotides comprising gene trapped sequences, useful in  
 PT functional genomic analysis, in the development of new therapeutic or  
 PT diagnostic agents, for diagnostic gene expression analysis or for  
 PT genetic manipulations  
 XX Claim 2; SEQ ID NO 1101; 29pp; English.  
 XX PS The present invention relates to novel murine cDNAs produced using  
 CC gene trap technology. The OMNITANK gene trapped sequences (GTSs)  
 CC are individually identified novel genes, and are useful in functional  
 CC genomic analysis, in the discovery and development of new therapeutic  
 CC and diagnostic agents, for gene discovery, for diagnostic gene  
 CC expression analysis, for cross species hybridisation analysis, and for  
 CC genetic manipulations such as antisense inhibition or gene targeting.

CC The polynucleotides of the invention are also useful for isolating  
 CC cDNAs, genomic clones or full-length genes/polynucleotides, or their  
 CC homologues, heterologues, paralogues or orthologues, that are capable  
 CC of hybridising to one or more of the new murine polynucleotide  
 CC sequences. The polynucleotides are also useful for identifying the  
 CC coding regions of the murine genome, and as hybridisation probes.  
 CC ABX90657-ABX91862 represent the murine GTSs of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC USPTO web site at [seqdata.uspto.gov/psipd/identity.html](http://seqdata.uspto.gov/psipd/identity.html).  
 XX SQ Sequence 516 BP; 163 A; 90 C; 118 G; 140 T; 5 other;  
 Query Match 82.0%; Score 16.4; DB 25; Length 516;  
 Best Local Similarity 94.4%; Pred. No. 1.3e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 GACAAAGCCCTGACAGCC 19  
 Db 100 GACAAAGCCCTGACAGCC 83  
 RESULT 3  
 ABZ74553  
 ID ABZ74553 standard; DNA; 8243 BP.  
 XX AC ABZ74553;  
 XX DT 12-MAY-2003 (first entry)  
 XX DE Secreted protein gene 346 genomic fragment HTHO21, SEQ ID NO:1700.  
 XX DE Human; secreted protein; cancer; tumour; hyperproliferative disorder;  
 KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;  
 KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;  
 KW drug screening; chromosome identification; chromosome mapping;  
 KW cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;  
 KW antineoplastic; vulnerrary; chromosome 15q24-25; gene; ds.  
 XX OS Homo sapiens.  
 XX PN WO200277013-A2.  
 XX PD 03-OCT-2002.  
 XX PF 26-MAR-2002; 2002WO-US09370.  
 XX PR 27-MAR-2001; 2001US-278650P.  
 XX PR 12-SEP-2001; 2001US-0950082.  
 XX PR 12-SEP-2001; 2001US-0950083.  
 XX PA (HUMA-) HUMAN GENOME SCI INC.  
 XX PI Rosen CA, Ruben SM;  
 XX WPI; 2003-040578/03.  
 XX DR WPI; 2003-040578/03.  
 XX PT New human secreted proteins and nucleic acids, useful for detecting or  
 PT treating cancer or other hyperproliferative disorders, autoimmune  
 PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia -  
 XX Disclosure; Page 2282-2284; 2474pp; English.  
 XX ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted  
 CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.  
 CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The  
 CC invention also encompasses antibodies specific for the secreted proteins,  
 CC the use of the secreted proteins in drug screening and recombinant  
 CC vectors and host cells comprising a nucleic acid of the invention. The  
 CC secreted proteins are thought to be involved in biological activities  
 CC associated with cellular signalling, cellular differentiation, cell  
 CC migration, prohormone activation and neurotransmitter activity. The  
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody

fragments specific for the secreted proteins, and modulators of protein activity are useful for diagnosing or treating cancers or other hyperproliferative disorders. Additionally, the secreted proteins and their nucleic acids may also be used in the treatment of autoimmune disorders, inflammatory disorders, diseases involving angiogenesis, AIDS (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote wound healing. Nucleic acids of the invention may be used for chromosome identification, chromosome mapping, in gene therapy, for identifying individuals from minute biological samples, as hybridisation probes, and as molecular weight markers. The present sequence represents a human secreted protein genomic fragment referred to in the disclosure of the invention.

Sequence 8243 BP; 2717 A; 1618 C; 1711 G; 2197 T; 0 other;

Query Match 82.0%; Score 16.4; DB 25; Length 8243;  
Best Local Similarity 94.4%; Pred. No. 1.9e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 GACAGCCTGACAGCC 19  
|||  
Db 8148 GACAGCCTGACAGCC 8165

## RESULT 4

AAK81012/C  
ID AAK81012 standard; DNA; 12710 BP.

AAK81012;

DT 07-NOV-2001 (first entry)

Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35824.

Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
cytostatic; gene therapy; vaccine; metastasis; ds.

Homo sapiens.

WO200157182-A2.

09-AUG-2001.

17-JAN-2001; 2001WO-US01354.

31-JAN-2000; 2000US-0179065.

04-FEB-2000; 2000US-0180628.

24-FEB-2000; 2000US-0184664.

02-MAR-2000; 2000US-0186350.

16-MAR-2000; 2000US-0189874.

17-MAR-2000; 2000US-0190076.

18-APR-2000; 2000US-0198123.

19-MAY-2000; 2000US-0205515.

07-JUN-2000; 2000US-0209467.

30-JUN-2000; 2000US-0214886.

07-JUL-2000; 2000US-0216880.

11-JUL-2000; 2000US-0217487.

11-JUL-2000; 2000US-0217496.

14-JUL-2000; 2000US-0218290.

26-JUL-2000; 2000US-0220963.

14-AUG-2000; 2000US-0224518.

14-AUG-2000; 2000US-0224519.

14-AUG-2000; 2000US-0225213.

14-AUG-2000; 2000US-0225214.

14-AUG-2000; 2000US-0225267.

14-AUG-2000; 2000US-0225268.

14-AUG-2000; 2000US-0225270.

14-AUG-2000; 2000US-0225447.

14-AUG-2000; 2000US-0225757.

14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231245.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0232968.  
PR 14-SEP-2000; 2000US-0233397.  
PR 14-SEP-2000; 2000US-0233398.  
PR 14-SEP-2000; 2000US-0233399.  
PR 14-SEP-2000; 2000US-0233400.  
PR 14-SEP-2000; 2000US-0233401.  
PR 14-SEP-2000; 2000US-0233402.  
PR 14-SEP-2000; 2000US-0233403.  
PR 14-SEP-2000; 2000US-0233404.  
PR 14-SEP-2000; 2000US-0233405.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234224.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235835.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239335.  
PR 13-OCT-2000; 2000US-0239337.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246532.

08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0251990.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 XX (HUMA-) HUMAN GENOME SCT INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI; 2001-483426/52.  
 XX  
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and  
 PS metastasis -  
 XX  
 XX  
 XX Disclosure; SEQ ID NO 35824; 3071bp + Sequence Listing; English.  
 XX  
 XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic  
 CC activity, and can be used in gene therapy and vaccine production. (II)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patients own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting  
 CC the nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/hematopoietic-related diseases, especially  
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/hematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
 CC represent sequences used in the exemplification of the present invention.  
 CC  
 XX  
 XX Sequence 12710 BP; 3496 A; 2567 C; 2575 G; 4072 T; 0 other;

Query Match 82.0%; Score 16.4; DB 22; Length 12710;  
 Best Local Similarity 94.4%; Pred. No. 2e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 GACAGCCTGACAGCC 19

Db 96 GACAGCCTGACAGCC 79

RESULT 5  
 ABZ74552  
 ID ABZ74552 strand; DNA; 12710 BP.  
 XX  
 XX ABZ74552;  
 AC  
 XX  
 XX 12-MAY-2003 (first entry)  
 DT  
 XX  
 XX  
 DE Secreted protein gene 346 genomic fragment HTHO21, SEQ ID NO:1699.  
 XX  
 XX  
 KW Human; secreted protein; cancer; tumour; hyperproliferative disorder;  
 KW autoimmune disorder; inflammation; angioinfectious diseases; AIDS;  
 KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;  
 KW drug screening; chromosome identification; chromosome mapping;  
 KW cytotoxic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;  
 KW antianaemic; vulnary; chromosome 15q24-25; gene; ds.  
 XX  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200277013-A2.  
 PN  
 XX  
 XX 03-OCT-2002.  
 PD  
 XX  
 XX 26-MAR-2002; 2002MO-US09370.  
 PF  
 XX  
 XX 27-MAR-2001; 2001US-278650P.  
 PR 12-SEP-2001; 2001US-0950082.  
 PR 12-SEP-2001; 2001US-0950083.  
 PR  
 XX  
 XX (HUMA-) HUMAN GENOME SCT INC.  
 PA  
 XX  
 XX Rosen CA, Ruben SM;  
 XX  
 XX WPI; 2003-040578/03.  
 DR  
 XX  
 XX  
 PT New human secreted proteins and nucleic acids, useful for detecting or  
 PT treating cancer or other hyperproliferative disorders, autoimmune  
 PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia -  
 PS  
 XX  
 XX Disclosure; Page 2278-2281; 2474bp; English.

ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted  
 CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.  
 CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The  
 CC invention also encompasses antibodies specific for the secreted proteins,  
 CC the use of the secreted proteins in drug screening and recombinant  
 CC vectors and host cells comprising a nucleic acid of the invention. The  
 CC secreted proteins are thought to be involved in biological activities  
 CC associated with cellular signalling, cellular differentiation, cell  
 CC migration, proinflammatory activation and neurotransmitter activity. The  
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody  
 CC fragments specific for the secreted proteins, and modulators of protein  
 CC activity are useful for diagnosing or treating cancers or other  
 CC hyperproliferative disorders. Additionally, the secreted proteins and  
 CC their nucleic acids may also be used in the treatment of autoimmune  
 CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS  
 CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote  
 CC wound healing. Nucleic acids of the invention may be used for chromosome  
 CC identification, chromosome mapping, in gene therapy, for identifying  
 CC individuals from minute biological samples, as hybridisation probes, and  
 CC as molecular weight markers. The present sequence represents a human  
 CC secreted protein genomic fragment referred to in the disclosure of the  
 CC invention.  
 CC  
 XX  
 XX Sequence 12710 BP; 4072 A; 2575 C; 2567 G; 3496 T; 0 other;

Query Match 82.0%; Score 16.4; DB 25; Length 12710;  
 Best Local Similarity 94.4%; Pred. No. 2e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;



QY 2 GACAGCCCTGACAGCC 19  
DB 12615 GACAGCCCTGAGAGCC 12632

RESULT 6  
ID ABX52814/c  
ABX52814 standard; cDNA; 433 BP.

AC ABX52814;

DT 25-FEB-2003 (first entry)

DE Bovine EST associated with lactation/muscle/fat deposition #2743.

XX Bovine; ss; EST; expressed sequence tag; lactation; LMPD;  
XX muscle deposition; fat deposition; genome mapping; gene identification;  
XX gene analysis; cattle breeding.

XX Bos Taurus.

XX US2002137160-A1.

XX 26-SEP-2002.

XX 26-OCT-2001; 2001US-0983965.

XX 17-DEC-1998; 98US-113678P.

XX 15-DEC-1999; 99US-0465231.

XX (BYAT/) BYAT J C.

XX (MATH/) MATHALAGAN N.

XX (TAON/) TAO N.

XX (WARR/) WARREN W C.

XX Byatt JC, Mathialagan N, Tao N, Warren WC;

XX WPI; 2003-102386/09.

XX Claim 2; SEQ ID NO 2743; 38pp; English.

XX The invention relates to a purified nucleic acid molecule associated with

XX lactation or muscle and fat deposition (designated LMPD), derived

XX from cattle, and the LMPD nucleic acid can specifically hybridise to a

XX second nucleic acid molecule comprising any of 5912 nucleotide

XX sequences, appearing as ABX50072-ABX55983, or complements of them.

XX Also included are: (1) a transformed cell having a nucleic acid

XX comprising an LMPD nucleic acid linked to a promoter and a 3' non-

XX translated sequence that functions in the cell to cause termination of

XX transcription and addition of polyadenylated ribonucleotides to a 3' end

XX of the mRNA molecule; and (2) determining a level or pattern of a

XX molecule in a bovine cell or tissue comprising: (a) incubating a marker

XX nucleic acid (comprising any of the 5912 nucleic acid sequences or its

XX complement or fragment) with a complementary nucleic acid molecule

XX obtained from the bovine cell or tissue, where hybridisation between the

XX marker nucleic acid and the complementary nucleic acid permits the

XX detection of the molecule; and (b) detecting the level or pattern of the

XX complementary nucleic acid, where the detection of the complementary

XX nucleic acid is predictive of the level or pattern of the molecule.

XX The LMPD nucleic acid is used for determining a level or pattern

XX of a molecule in a bovine cell or tissue. It is useful for genome

XX mapping, gene identification and analysis, cattle breeding, preparation

XX of constructs for use in cattle gene expression, or for genetically

XX improving cattle. The present sequence is one of the 5912 bovine

XX LMPD EST (expressed sequence tag) nucleic acids.

XX Note: The present sequence was not shown in the specification but

XX was obtained in electronic format from the USPTO web site:

XX seqdata.uspto.gov/sequence.html?DocID=20020137160.

XX SQ Sequence 433 BP; 86 A; 88 C; 159 G; 100 T; 0 other;  
Query Match 80.0%; Score 16; DB 25; Length 433;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CAAGCCCTGACAGCC 19  
DB 102 CAAGCCCTGACAGCC 87

RESULT 7  
ID ABX27459  
ABX27459 standard; cDNA; 165 BP.

XX AC ABX27459;

XX DT 11-FEB-2003 (first entry)

XX DE Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #9516.

XX Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;

XX cellular fucosylation; glycoconjugate fucosylation; transplant rejection;

XX arthritis; asthma; sepsis; reperfusion injury; stroke; infection;

XX complex carbohydrate; gene replacement therapy; immunosuppressive;

XX antiinflammatory; antiarthritic; antibacterial; cerebroprotective;

XX antiaesthetic; vasotropic.

XX Homo sapiens.

XX US2002110548-A1.

XX 15-AUG-2002.

XX 11-JUN-2001; 2001US-0878574.

XX 22-NOV-1996; 96US-0753233.

XX 03-DEC-1997; 97US-0984246.

XX 09-SEP-1998; 98US-0149674.

XX 14-JUN-1999; 99US-0333177.

XX (GENY ) GENETICS INST INC.

XX Sullivan F, Kriz R, Kumar R;

XX WPI; 2003-066673/06.

XX New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)

XX peptide, for manufacturing complex carbohydrates, or as targets for

XX screening GM4,6D antagonists for treating e.g. arthritis, or transplant

XX rejection

XX Disclosure; SEQ ID NO 9518; 6pp; English.

XX The invention relates to a composition comprising a human GDP-mannose.

XX 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying

XX GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation

XX in a mammalian subject and for treating or ameliorating diseases affected

XX by the level of cellular fucosylation or diseases affected by the

XX fucosylation of glycoconjugates. These diseases include arthritis,

XX transplant rejection, asthma, sepsis, reperfusion injury, stroke or

XX infection. The GM4,6D peptide or a polynucleotide encoding it is also

XX useful for manufacturing complex carbohydrates and as targets for

XX screening small molecule antagonists of the activity of the enzyme. The

XX polynucleotide is useful in developing an assay for defects in the

XX enzyme, as well as in gene replacement therapy. Sequences

XX ABX17942-ABX17944 and ABX17947-ABX33716 represent DNA molecules encoding

XX human GM4,6D peptides of the invention.

XX Note: The sequence data for this patent did not form part of the printed

XX specification but was obtained in electronic format directly from USPTO

XX at seqdata.uspto.gov/sequence.html.

SO Sequence 165 BP; 34 A; 66 C; 42 G; 23 T; 0 other;  
Query Match 79.0%; Score 15.8; DB 25; Length 165;  
Best Local Similarity 89.5%; Pred. No. 2.3e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GGACAGCCCTGACAGCCA 19  
DB 131 GGACAGCCCTGACAGCCA 149

RESULT 8  
ABZ41531  
ID ABZ41531 standard; DNA; 261 BP.  
XX  
AC ABZ41531;  
XX  
DT 07-MAR-2003 (first entry)  
XX  
DE N. gonorrhoeae nucleotide sequence SEQ ID 7651.  
XX  
KM Antibacterial; infection; vaccine; gene therapy; gene; ds.  
XX  
OS Neisseria gonorrhoeae.  
XX  
PN WO200279243-A2.  
XX  
PD 10-OCT-2002.  
XX  
PF 12-FEB-2002; 2002WO-IB02069.  
XX  
PR 12-FEB-2001; 2001GB-0003424.  
XX  
PA (CHIR-) CHIRON SPA.  
XX  
PI Fontana MR, Pizza M, Maignani V, Monaci E;  
XX  
DR WPI; 2003-058415/05.  
XX  
P-PSDB; ABP80561.  
XX  
PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a  
XX  
PT medicament for treating or preventing N. gonorrhoeae infection -  
XX  
PS Disclosure; Page 746; 815pp; English.  
XX  
CC The present invention relates to proteins from Neisseria gonorrhoeae.  
XX  
CC Also disclosed are the nucleic acid molecules encoding the proteins and  
XX  
CC antibodies that specifically bind to the proteins. The composition  
XX  
CC comprising the protein, nucleic acid or antibody is useful for the  
XX  
CC manufacture of a medicament for treating or preventing N. gonorrhoeae  
XX  
CC infection, this may be in the form of a vaccine or gene therapy.  
XX  
CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid  
XX  
CC molecules of the invention.  
XX  
SQ Sequence 261 BP; 53 A; 71 C; 65 G; 72 T; 0 other;  
Query Match 79.0%; Score 15.8; DB 25; Length 261;  
Best Local Similarity 89.5%; Pred. No. 2.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GACAGCCCTGACAGCCA 20  
DB 153 GACAGCCCTGACAGCCA 171

RESULT 9  
AAZ11974/c  
ID AAZ11974 standard; DNA; 273 BP.  
XX  
AC AAZ11974;  
XX  
DT 08-OCT-1999 (first entry)  
XX

DE Neisseria meningitidis strain A complete ORF37 sequence.  
XX  
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
XX  
KW treatment; Neisseria infection; meningitis; septicemia; gonorrhea; ss.  
XX  
OS Neisseria meningitidis.  
XX  
PN WO9924578-A2.  
XX  
PD 20-MAY-1999.  
XX  
PF 09-OCT-1998; 98WO-IB01665.  
XX  
PR 01-SEP-1998; 98GB-0019016.  
XX  
PR 06-NOV-1997; 97GB-0023516.  
XX  
PR 14-NOV-1997; 97GB-0024190.  
XX  
PR 18-NOV-1997; 97GB-0024386.  
XX  
PR 27-NOV-1997; 97GB-0025158.  
XX  
PR 10-DEC-1997; 97GB-0026147.  
XX  
PR 14-JUN-1998; 98GB-0000759.  
XX  
PA (CHIR-) CHIRON SPA.  
XX  
PI Grandi G, Maignani V, Pizza M, Rappuoli R, Scarlato V;  
XX  
XX  
DR WPI; 1999-327407/27.  
XX  
P-PSDB; AAY38501.  
XX  
PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for  
XX  
PT diagnosis, treatment and prevention of infection  
XX  
PS Claim 3; Page 61; 524pp; English.  
XX  
CC Nucleotide sequences AAZ11972-Z12358 represent open reading frames  
XX  
CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode  
XX  
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their  
XX  
CC fragments, their nucleic acids and antibodies are used for diagnosis,  
XX  
CC prevention (as vaccines) or treatment of Neisseria infections,  
XX  
CC such as meningitis, septicemia and gonorrhea. Both organisms  
XX  
CC are closely related. Fragments of the nucleic acids are useful  
XX  
CC as hybridisation probes and antisense reagents.  
XX  
SQ Sequence 273 BP; 75 A; 67 C; 83 G; 48 T; 0 other;  
Query Match 79.0%; Score 15.8; DB 20; Length 273;  
Best Local Similarity 89.5%; Pred. No. 2.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GACAGCCCTGACAGCCA 20  
DB 217 GACAGCCCTGACAGCCA 199

RESULT 10  
AAZ53178  
ID AAZ53178 standard; DNA; 288 BP.  
XX  
AC AAZ53178;  
XX  
DT 21-MAR-2000 (first entry)  
XX  
DE Neisseria meningitidis ORF 091 partial DNA sequence SEQ ID NO:307.  
XX  
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
XX  
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
XX  
KW antibacterial; gene therapy; ds.  
XX  
OS Neisseria meningitidis.  
XX  
PN WO9957280-A2.  
XX  
PD 11-NOV-1999.  
XX

PF 30-APR-1999; 99WO-US09346.  
 XX 01-MAY-1998; 98US-0083758.  
 PR 31-JUL-1998; 98US-0094869.  
 PR 02-SEP-1998; 98US-0098994.  
 PR 02-SEP-1998; 98US-0099062.  
 PR 09-OCT-1998; 98US-0103749.  
 PR 09-OCT-1998; 98US-0103794.  
 PR 09-OCT-1998; 98US-0103796.  
 PR 25-FEB-1999; 99US-0121528.  
 XX (CHIR) CHIRON CORP.  
 PA (GENO-) INST GENOMIC RES.  
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;  
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
 PI Tettelin H, Venter JC;  
 DR WPI; 2000-062150/05.  
 DR P-PSDB; AAY74416.  
 PT Novel Neisserial polypeptides predicted to be useful antigens for  
 PT vaccines and diagnostics  
 PS Claim 7; Page 287; 1453pp; English.  
 XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941  
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides  
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent  
 CC PCR primers used in the exemplification of the present invention. The  
 CC polypeptides, the polynucleotides, antibodies and compositions of  
 CC the invention can be used as vaccines, as diagnostic reagents, and as  
 CC immunogenic compositions. The polypeptides can be used in the  
 CC manufacture of medicaments for treating or preventing infection due to  
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the  
 CC presence of Neisseria bacteria, or to raise antibodies. They may also  
 CC be used to screen for agonists or antagonists, which may themselves  
 CC have use as antibacterial agents. The polynucleotides of the invention  
 CC may also be used in gene therapy protocols.  
 XX  
 SQ Sequence 288 BP; 57 A; 73 C; 74 G; 84 T; 0 other;  
 Query Match 79.0%; Score 15.8; DB 21; Length 288;  
 Best Local Similarity 89.5%; Pred. No. 2.4e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GACAAAGCCCTGACAGCCA 20  
 DB 198 GACAAAGCCCTGACAGCCA 216  
 RESULT 11  
 AAZ53176  
 ID AAZ53176 standard; DNA; 309 BP.  
 XX  
 XX AAZ53176;  
 DT 21-MAR-2000 (first entry)  
 XX  
 DE Neisseria gonorrhoeae ORF 091 partial DNA sequence SEQ ID NO:303.  
 XX  
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
 KM antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
 KM antibacterial; gene therapy; ds.  
 XX  
 XX Neisseria gonorrhoeae.  
 OS  
 XX  
 PN WO9957280-A2.  
 XX  
 PD 11-NOV-1999.  
 XX  
 PF 30-APR-1999; 99WO-US09346.  
 XX

PR 01-MAY-1998; 98US-0083758.  
 PR 31-JUL-1998; 98US-0094869.  
 PR 02-SEP-1998; 98US-0098994.  
 PR 02-SEP-1998; 98US-0099062.  
 PR 09-OCT-1998; 98US-0103749.  
 PR 09-OCT-1998; 98US-0103794.  
 PR 09-OCT-1998; 98US-0103796.  
 PR 25-FEB-1999; 99US-0121528.  
 XX (CHIR) CHIRON CORP.  
 PA (GENO-) INST GENOMIC RES.  
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;  
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
 PI Tettelin H, Venter JC;  
 DR WPI; 2000-062150/05.  
 DR P-PSDB; AAY74416.  
 PT Novel Neisserial polypeptides predicted to be useful antigens for  
 PT vaccines and diagnostics  
 PS Claim 7; Page 287; 1453pp; English.  
 XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941  
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides  
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent  
 CC PCR primers used in the exemplification of the present invention. The  
 CC polypeptides, the polynucleotides, antibodies and compositions of  
 CC the invention can be used as vaccines, as diagnostic reagents, and as  
 CC immunogenic compositions. The polypeptides can be used in the  
 CC manufacture of medicaments for treating or preventing infection due to  
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the  
 CC presence of Neisseria bacteria, or to raise antibodies. They may also  
 CC be used to screen for agonists or antagonists, which may themselves  
 CC have use as antibacterial agents. The polynucleotides of the invention  
 CC may also be used in gene therapy protocols.  
 XX  
 SQ Sequence 309 BP; 64 A; 82 C; 78 G; 85 T; 0 other;  
 Query Match 79.0%; Score 15.8; DB 21; Length 309;  
 Best Local Similarity 89.5%; Pred. No. 2.4e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GACAAAGCCCTGACAGCCA 20  
 DB 198 GACAAAGCCCTGACAGCCA 216  
 RESULT 12  
 AAZ1975/c  
 ID AAZ1975 standard; DNA; 381 BP.  
 XX  
 XX AAZ1975;  
 DT 08-OCT-1999 (first entry)  
 XX  
 DE Neisseria gonorrhoeae complete ORF37 sequence.  
 XX  
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
 KM treatment; Neisseria infection; meningitis; septicemia; gonorrhea; ss.  
 OS  
 XX Neisseria gonorrhoeae.  
 PN WO9924578-A2.  
 XX  
 PD 20-MAY-1999.  
 XX  
 PF 09-OCT-1998; 98WO-IB01665.  
 XX  
 PR 01-SEP-1998; 98GB-0019016.  
 PR 06-NOV-1997; 97GB-0023516.  
 PR 14-NOV-1997; 97GB-0024190.  
 PR

PR 18-NOV-1997; 97GB-0024386.  
PR 27-NOV-1997; 97GB-0025158.  
PR 10-DEC-1997; 97GB-0026147.  
PR 14-JAN-1998; 98GB-0000759.  
XX

PA (CHIR-) CHIRON SPA.

XX Grandi G, Masiagnani V, Pizzia M, Rappuoli R, Scarlato V,  
XX P-PSDB; AAY38502.

DR WPI; 1999-327407/27.

XX P-PSDB; AAY38502.

PT Proteins from *Neisseria meningitidis* and *N. gonorrhoeae* useful for  
diagnosis, treatment and prevention of infection

XX Claim 3; Page 61; 524pp; English.

CC Nucleotide sequences AA211972-212358 represent open reading frames  
(ORFs) of *Neisseria meningitidis* and *N. gonorrhoeae* which encode  
antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their  
fragments, their nucleic acids and antibodies are used for diagnosis, their  
prevention (as vaccines) or treatment of *Neisseria* infections,  
such as meningitis, septicemia and gonorrhea. Both organisms  
are closely related. Fragments of the nucleic acids are useful  
as hybridisation probes and antisense reagents.

SO Sequence 381 BP; 106 A; 84 C; 114 G; 77 T; 0 other;

Query Match 79.0%; Score 15.8; DB 20; Length 381;  
Best Local Similarity 89.5%; Pred. No. 2.5e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GACAGCCCTGACAGCCA 20  
|||

Db 325 GACAGCCCTGACAGCCA 307  
|||

RESULT 13

AAKS8992 ID AAKS8992 standard; cDNA; 409 BP.

XX AAKS8992;

DT 06-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:4052.

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
cytotoxic; gene therapy; vaccine; metastasis; ss.

OS Homo sapiens.

FN WO200157182-A2.

XX 09-AUG-2001.

PF 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0188874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 28-JUN-2000; 2000US-0209467.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216847.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 14-AUG-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0225759.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 06-SEP-2000; 2000US-0230437.

PR 08-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231411.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0232080.

PR 12-SEP-2000; 2000US-0232081.

PR 14-SEP-2000; 2000US-0232397.

PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 21-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 25-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.

PR 26-SEP-2000; 2000US-0234998.

PR 27-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.

PR 02-OCT-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.

PR 13-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239335.

PR 20-OCT-2000; 2000US-0239337.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.

PR 20-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.



PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 12-SEP-2000; 2000US-0232081.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
  
(HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-483426/52.  
XX  
DR  
XX  
XX  
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
XX metastasis -  
XX  
PS  
XX  
XX  
CC Disclosure; SEQ ID NO 29054; 3071pp + Sequence Listing; English.  
CC  
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 749 BP; 201 A; 185 C; 137 G; 226 T; 0 other;

Query Match 79.0%; Score 15.8; DB 22; Length 749;  
Best Local Similarity 89.5%; Pred. No. 2.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCACAAAGCCCTGACAAAGCC 19  
DB 558 GCACATGCCCGACAAAGCC 576

RESULT 15  
AAK74243

ID AAK74243 standard; DNA; 749 BP.  
XX AAK74243;  
XX  
XX 07-NOV-2001 (first entry)  
XX  
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29055.  
DE Human immune/haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytoskeletal; gene therapy; vaccine; metastasis; ds  
XX Homo sapiens.  
OS  
XX MO200157182-A2.  
PN  
XX 09-AUG-2001.  
PD  
XX 17-JAN-2001; 2001WO-US01354.  
PF  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214686.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.

PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259676.  
 XX

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

DR WPI; 2001-483426/52.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and  
 PT metastasis -

PS Disclosure; SEQ ID NO 29055; 3071pp + Sequence Listing; English.

CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patients own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting  
 CC the nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/haematopoietic-related diseases, especially  
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/haematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
 CC represent sequences used in the exemplification of the present invention.  
 XX

XX Sequence 749 BP; 201 A; 185 C; 137 G; 226 T; 0 other;

Query Match 79.0%; Score 15.8; DB 22; Length 749;

Best Local Similarity 89.5%; Pred. No. 2.7e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCC 19

DB 558 GGACATGCCCGACAGCC 576

Search completed: August 1, 2003, 13:55:23  
 Job time : 6.69868 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 08:05:17 ; Search time 5.18607 Seconds  
(without alignments)  
7955.924 Million cell updates/sec

Title: US-09-834-291-24

Perfect score: 1 ggacagccctgacagcca 20

Sequence: IDENTITY\_NUC

Scoring table: Gapop 10.0, Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues 2879534

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/PCUS\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09D\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	20	9	US-09-834-291-5	Sequence 5, Appli
2	20	100.0	20	9	US-09-834-291-10	Sequence 10, Appli
3	20	100.0	20	9	US-09-834-291-24	Sequence 24, Appli
4	20	100.0	266	9	US-09-834-291-32	Sequence 32, Appli
5	20	100.0	720	9	US-09-834-291-2	Sequence 2, Appli
6	20	100.0	2380	9	US-09-834-291-3	Sequence 3, Appli
7	20	100.0	2827	9	US-09-834-291-4	Sequence 4, Appli
8	20	100.0	3212	9	US-09-834-291-1	Sequence 1, Appli
9	18.4	92.0	20	9	US-09-834-291-6	Sequence 6, Appli
10	16.8	84.0	20	9	US-09-834-291-7	Sequence 7, Appli
11	16.8	84.0	20	9	US-09-834-291-9	Sequence 9, Appli
12	16.8	84.0	2940917	13	US-10-027-632-174763	Sequence 174763, Appli
13	16.4	82.0	516	10	US-09-728-444-1101	Sequence 1101, Ap
14	16.4	80.0	433	10	US-09-834-291-3	Sequence 2743, Ap
15	15.8	79.0	165	10	US-09-878-574-9518	Sequence 9518, Ap
16	15.8	79.0	487	13	US-10-027-632-87503	Sequence 87503, A

17	15.8	79.0	487	13	US-10-027-632-316681	Sequence 316681, Appli
18	15.8	79.0	88421	10	US-09-976-059-1	Sequence 1, Appli
19	15.8	79.0	155074	13	US-10-026-188-6	Sequence 6, Appli
20	15.8	79.0	3309400	10	US-09-738-626-1	Sequence 1, Appli
21	15.8	79.0	9025608	14	US-10-156-761-1	Sequence 1975, Ap
22	15.4	77.0	272	10	US-09-983-965-3975	Sequence 493, App
23	15.4	77.0	401	9	US-09-795-686-493	Sequence 493, App
24	15.4	77.0	401	9	US-09-795-686-493	Sequence 493, App
25	15.4	77.0	401	9	US-09-795-686-493	Sequence 493, App
26	15.4	77.0	664	13	US-10-027-632-263366	Sequence 966, App
27	15.4	77.0	673	10	US-09-764-877-966	Sequence 966, App
28	15.4	77.0	845	10	US-09-822-830A-557	Sequence 966, App
29	15.4	77.0	860	10	US-09-764-877-3925	Sequence 3925, Ap
30	15.4	77.0	860	10	US-09-764-877-3926	Sequence 3926, Ap
31	15.4	77.0	872	13	US-10-027-632-127733	Sequence 127733, Ap
32	15.4	77.0	872	13	US-10-027-632-127734	Sequence 127734, Ap
33	15.4	77.0	2155	13	US-10-027-632-102861	Sequence 102861, Ap
34	15.4	77.0	2155	13	US-10-027-632-102862	Sequence 102862, Ap
35	15.4	77.0	3038	14	US-10-128-714-2538	Sequence 258, App
36	15.4	77.0	3604	14	US-10-128-714-2538	Sequence 258, App
37	15.4	77.0	34668	11	US-09-900-449A-3	Sequence 5258, Ap
38	15.4	77.0	1503841	9	US-09-795-686-1	Sequence 1, Appli
39	15.4	77.0	1503841	9	US-09-795-686-1	Sequence 1, Appli
40	15.4	77.0	424	10	US-09-946-807-1	Sequence 1862, Ap
41	15.2	76.0	424	10	US-09-983-965-2134	Sequence 2134, Ap
42	15.2	76.0	428	10	US-09-883-965-2134	Sequence 2134, Ap
43	15.2	76.0	502	10	US-09-783-590-2161	Sequence 2161, Ap
44	15.2	76.0	663	13	US-10-027-632-187538	Sequence 187538, Ap
45	15.2	76.0	663	13	US-10-027-632-187539	Sequence 187539, Ap

## ALIGNMENTS

RESULT 1  
US-09-834-291-5  
Sequence 5, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-132  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-5

Query Match 100.0%; Score 20; DB 9; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.2; Mismatches 0; Indels 0; Gaps 0;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20  
DB 1 GGACAGCCCTGACAGCCA 20

RESULT 2  
US-09-834-291-10  
Sequence 10, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
TITLE OF INVENTION: p53 Binding Areas  
FILE REFERENCE: 4121-132  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-5

APPLICANT: Muller-Schilling, Martina  
; TITLE OF INVENTION: p53 Binding Areas  
; FILE REFERENCE: 4121-122  
; CURRENT APPLICATION NUMBER: US/09/834,291  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/DE99/03343  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: DE 198 47 779.1  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-09-834-291-10

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 20;  
Matches 20; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAGCCA 20  
DB 1 GGACAAAGCCCTGACAGCCA 20

RESULT 3  
US-09-834-291-24  
; Sequence 24, Application US/09834291  
; Patent No. US20020042064A1  
; GENERAL INFORMATION:  
; APPLICANT: Kramer, Peter  
; APPLICANT: Muller-Schilling, Martina  
; TITLE OF INVENTION: p53 Binding Areas  
; FILE REFERENCE: 4121-122  
; CURRENT APPLICATION NUMBER: US/09/834,291  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/DE99/03343  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: DE 198 47 779.1  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 24  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-09-834-291-24

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 20;  
Matches 20; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAGCCA 20  
DB 1 GGACAAAGCCCTGACAGCCA 20

RESULT 4  
US-09-834-291-32  
; Sequence 32, Application US/09834291  
; Patent No. US20020042064A1  
; GENERAL INFORMATION:  
; APPLICANT: Kramer, Peter  
; APPLICANT: Muller-Schilling, Martina  
; APPLICANT: Oren, Moshe  
; TITLE OF INVENTION: p53 Binding Areas  
; FILE REFERENCE: 4121-122  
; CURRENT APPLICATION NUMBER: US/09/834,291  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/DE99/03343

PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: DE 198 47 779.1  
; PRIOR FILING DATE: 1998-10-16  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 32  
; LENGTH: 266  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-09-834-291-32

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 266;  
Matches 20; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAGCCA 20  
DB 160 GGACAAAGCCCTGACAGCCA 179

RESULT 5  
US-09-834-291-2  
; Sequence 2, Application US/09834291  
; Patent No. US20020042064A1  
; GENERAL INFORMATION:  
; APPLICANT: Kramer, Peter  
; APPLICANT: Muller-Schilling, Martina  
; APPLICANT: Oren, Moshe  
; TITLE OF INVENTION: p53 Binding Areas  
; FILE REFERENCE: 4121-122  
; CURRENT APPLICATION NUMBER: US/09/834,291  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/DE99/03343  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: DE 198 47 779.1  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 720  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-09-834-291-2

Query Match  
Best Local Similarity 100.0%; Score 20; DB 9; Length 720;  
Matches 20; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAGCCA 20  
DB 160 GGACAAAGCCCTGACAGCCA 179

RESULT 6  
US-09-834-291-3  
; Sequence 3, Application US/09834291  
; Patent No. US20020042064A1  
; GENERAL INFORMATION:  
; APPLICANT: Kramer, Peter  
; APPLICANT: Muller-Schilling, Martina  
; APPLICANT: Oren, Moshe  
; TITLE OF INVENTION: p53 Binding Areas  
; FILE REFERENCE: 4121-122  
; CURRENT APPLICATION NUMBER: US/09/834,291  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/DE99/03343  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: DE 198 47 779.1  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2380

TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-3

Query Match 100.0%; Score 20; DB 9; Length 2380;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20  
DB 1820 GGACAGCCCTGACAGCCA 1839

RESULT 7  
US-09-834-291-4  
Sequence 4, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2827  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-4

Query Match 100.0%; Score 20; DB 9; Length 2827;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20  
DB 2267 GGACAGCCCTGACAGCCA 2266

RESULT 8  
US-09-834-291-1  
Sequence 1, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 1  
LENGTH: 3212  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-1

Query Match 100.0%; Score 20; DB 9; Length 3212;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20  
DB 2659 GGACAGCCCTGACAGCCA 2678

RESULT 9  
US-09-834-291-6  
Sequence 6, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 6  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-6

Query Match 92.0%; Score 18.4; DB 9; Length 20;  
Best Local Similarity 95.0%; Pred. No. 7.4;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20  
DB 1 GGAAAGCCCTGACAGCCA 20

RESULT 10  
US-09-834-291-7  
Sequence 7, Application US/09834291  
Patent No. US20020042064A1  
GENERAL INFORMATION:  
APPLICANT: Krammer, Peter  
APPLICANT: Muller-Schilling, Martina  
APPLICANT: Oren, Moshe  
FILE REFERENCE: 4121-122  
CURRENT APPLICATION NUMBER: US/09/834,291  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/DE99/03343  
PRIOR FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: DE 198 47 779.1  
PRIOR FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 7  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-09-834-291-7

Query Match 84.0%; Score 16.8; DB 9; Length 20;  
Best Local Similarity 90.0%; Pred. No. 46;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20  
DB 1 GGAAAGCCCTGAAAGCCA 20

RESULT 11

US-09-834-291-9  
; Sequence 9, Application US/09834291  
; Patent No. US20020042064A1  
; GENERAL INFORMATION:  
; APPLICANT: Krammer, Peter  
; APPLICANT: Muller-Schilling, Martina  
; APPLICANT: Oren, Moshe  
; TITLE OF INVENTION: p53 Binding Areas  
; FILE REFERENCE: 4121-122  
; CURRENT APPLICATION NUMBER: US/09/834,291  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/DE99/03343  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: DE 198 47 779.1  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-09-834-291-9

Query Match 84.0%; Score 16.8; DB 9; Length 20;  
Best Local Similarity 90.0%; Pred. No. 46;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GCACAGCCCTGACAGCCA 20  
DB 1 GCACAGCCCTGACAGCCA 20

RESULT 12  
US-10-027-632-174763/C  
; Sequence 174763, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 174763  
; LENGTH: 2940917  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(2940917)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-027-632-174763

Query Match 84.0%; Score 16.8; DB 13; Length 2940917;  
Best Local Similarity 90.0%; Pred. No. 46;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GCACAGCCCTGACAGCCA 20

DB 2746843 GCACAGCCCTGAATGCCA 2746824

RESULT 13  
US-09-728-444-1101/C  
; Sequence 1101, Application US/09728444  
; Patent No. US20020161207A1  
; GENERAL INFORMATION:  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sambrowicz, Brian  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: No. US20020161207A1 Murine Polynucleotide Sequences  
; FILE REFERENCE: LEX-0100-USA  
; CURRENT APPLICATION NUMBER: US/09/728,444  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/168,360  
; PRIOR FILING DATE: 1999-12-01  
; NUMBER OF SEQ ID NOS: 1206  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1101  
; LENGTH: 516  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(516)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-728-444-1101

Query Match 82.0%; Score 16.4; DB 10; Length 516;  
Best Local Similarity 94.4%; Pred. No. 74;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GCACAGCCCTGACAGCC 19  
DB 100 GCACAGCCCTGACAGCC 83

RESULT 14  
US-09-983-965-2743/C  
; Sequence 2743, Application US/09983965  
; Patent No. US20020137160A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengding  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 37-21 (10297)C  
; CURRENT APPLICATION NUMBER: US/09/983,965  
; PRIOR FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: US 09/465,231  
; PRIOR FILING DATE: 1999-12-15  
; PRIOR APPLICATION NUMBER: US 60/113,678  
; PRIOR FILING DATE: 1998-12-17  
; NUMBER OF SEQ ID NOS: 5912  
; SEQ ID NO 2743  
; LENGTH: 433  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 13-LIB3058-025-Q1-K1-D1  
US-09-983-965-2743

Query Match 80.0%; Score 16; DB 10; Length 433;  
Best Local Similarity 100.0%; Pred. No. 1,2e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 CAAGCCCTGACAGCC 19

Db \* 102 CAAGCCCTGACAAGCC 87

## RESULT 15

US-09-878-574-9518  
; Sequence 9518, Application US/09878574  
; Patent No. US20020110548A1

## GENERAL INFORMATION:

; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 9518  
; LENGTH: 165  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: 701102580H1  
US-09-878-574-9518

Query Match 79.0%; Score 15.8; DB 10; Length 165;  
Best Local Similarity 89.5%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGACAGCCCTGACAAGCC 19

Db 131 GGACAGCTCGACAAGCC 149

Search completed: August 1, 2003, 13:37:14  
Job time: 17.1861 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 02:21:37 ; Search time 1.65066 Seconds  
(without alignments)  
5347.959 Million cell updates/sec

Title: US-09-834-291-24  
Perfect score: 20  
Sequence: 1 ggaacagccctgacacgcca 20

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 569978 segs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/prodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/prodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/prodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/prodata/1/ina/PCUS.COMB.seq.\*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15.2	76.0	1923	US-09-220-132-85	Sequence 85, Appl
2	15.2	76.0	3554	US-08-460-309-1	Sequence 1, Appl
3	15.2	76.0	3554	US-08-125-077-1	Sequence 1, Appl
4	15.2	76.0	3554	US-08-460-309-1	Sequence 1, Appl
5	15.2	76.0	6942	US-08-460-309-3	Sequence 3, Appl
6	15.2	76.0	6942	US-08-125-077-3	Sequence 3, Appl
7	14.8	74.0	1185	US-09-252-991A-5635	Sequence 5635, Ap
8	14.8	74.0	2214	US-09-489-847-113	Sequence 113, App
9	14.4	72.0	360	US-09-252-991A-8300	Sequence 8300, Ap
10	14.4	72.0	417	US-09-252-991A-8306	Sequence 8306, Ap
11	14.4	72.0	1237	US-08-240-372-1	Sequence 1, Appl
12	14.4	72.0	2470	US-08-564-805-221	Sequence 221, App
13	14.4	72.0	6312	US-09-011-745-7	Sequence 7, Appl
14	14.4	72.0	8088	US-09-315-127-4	Sequence 4, Appl
15	14.4	72.0	8535	US-08-716-351A-1	Sequence 1, Appl
16	14.4	72.0	36181	US-08-311-731A-120	Sequence 120, App
17	14.4	71.0	20	US-08-392-542-35	Sequence 35, Appl
18	14.2	71.0	20	US-08-894-327-35	Sequence 35, Appl
19	14.2	71.0	20	US-09-685-027-35	Sequence 35, Appl
20	14.2	71.0	30	US-08-347-792-18	Sequence 18, Appl
21	14.2	71.0	30	US-08-311-357-18	Sequence 29, Appl
22	14.2	71.0	30	US-08-697-221-29	Sequence 31, Appl
23	14.2	71.0	30	US-08-392-542-31	Sequence 31, Appl
24	14.2	71.0	30	US-08-894-327-31	Sequence 31, Appl
25	14.2	71.0	30	US-09-685-027-31	Sequence 31, Appl
26	14.2	71.0	30	PCT-US95-15153-18	Sequence 18, Appl
27	14.2	71.0	34	US-08-713-052-5	Sequence 5, Appl

28	14.2	71.0	72	US-09-276-531-93	Sequence 93, Appl
29	14.2	71.0	218	US-09-188-930-86	Sequence 86, Appl
30	14.2	71.0	218	US-09-312-283C-86	Sequence 3, Appl
31	14.2	71.0	538	US-08-979-586-3	Sequence 3, Appl
32	14.2	71.0	538	US-09-577-640-3	Sequence 28, Appl
33	14.2	71.0	780	US-08-445-909A-28	Sequence 11, Appl
34	14.2	71.0	1029	US-09-216-393B-11	Sequence 10642, A
35	14.2	71.0	1254	US-09-252-991A-10432	Sequence 68, Appl
36	14.2	71.0	1568	US-09-247-155-68	Sequence 10432, A
37	14.2	71.0	2169	US-09-252-991A-10524	Sequence 10524, A
38	14.2	71.0	2979	US-09-220-132-16	Sequence 16, Appl
39	14.2	71.0	3183	US-09-620-312D-571	Sequence 571, App
40	14.2	71.0	3470	US-09-894-997-49	Sequence 49, Appl
41	14.2	71.0	4057	US-08-726-214-5	Sequence 5, Appl
42	14.2	71.0	4533	US-08-726-214-5	Sequence 3, Appl
43	14.2	71.0	6942	US-08-460-309-3	Sequence 3, Appl
44	14.2	71.0	6942	US-08-125-077-3	Sequence 3, Appl
45	14.2	71.0	1230025	US-09-198-452A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1  
US-09-220-132-85  
Sequence 85, Application US/09220132  
Patent No. 650607  
GENERAL INFORMATION:  
APPLICANT: Shvian, Andrew M.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT  
FILE REFERENCE: 07334-074001  
CURRENT FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: US 60/079,303  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: US 60/068,821  
PRIOR FILING DATE: 1997-12-24  
NUMBER OF SEQ ID NOS: 191  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 85  
LENGTH: 1923  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-220-132-85  
Query Match 76.0%; Score 15.2; DB 4; Length 1923;  
Best Local Similarity 85.0%; Pred. No. 77;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 GCACAGCCCTGACACGCA 20  
Db 1572 GCAAGCCCTGACACGCA 1591  
RESULT 2  
US-08-460-309-1  
Sequence 1, Application US/08460309  
Patent No. 5837496  
GENERAL INFORMATION:  
APPLICANT: Enqvall, Eva  
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
TITLE OF INVENTION: Fragments and Uses Thereof  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,309
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/125,077
FILING DATE: 22-SEP-1993
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/919,951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3554 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 10..3400
US-08-460-309-1

Query Match
Best Local Similarity 76.0%; Score 15.2; DB 2; Length 3554;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20
183 GGACAAAGCCGACAAAGCCA 202

RESULT 3
US-08-125-077-1
Sequence 1, Application US/08125077
Patent No. 5872231
Patent No. 5872231 5840863
GENERAL INFORMATION:
APPLICANT: Engvall, Eva
APPLICANT: Leivo, Ilmo
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
NUMBER OF INVENTION: Fragments and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/125,077
FILING DATE: 22-SEP-1993
CLASSIFICATION: 435
```

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/919,951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3554 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 10..3400
US-08-125-077-1

Query Match
Best Local Similarity 76.0%; Score 15.2; DB 2; Length 3554;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```

QY 1 GGACAAAGCCCTGACAAAGCCA 20
183 GGACAAAGCCGACAAAGCCA 202
```

```

RESULT 4
5444158-1
APPLICANT: ENGVAL, EVA; SANES, JOSHUA
TITLE OF INVENTION: MEROSIN, NUCLEIC ACIDS ENCODING,
FRAGMENTS AND USES THEREOF
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/87,642
FILING DATE: 08-JUL-1993
APPLICATION NUMBER: 587,689
FILING DATE: 24-SEP-1990
APPLICATION NUMBER: 472,319
FILING DATE: 30-JAN-1990
SEQ ID NO: 1
LENGTH: 3554
5444158-1
```

```

Query Match
Best Local Similarity 76.0%; Score 15.2; DB 6; Length 3554;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20
183 GGACAAAGCCGACAAAGCCA 202
```

```

RESULT 5
US-08-460-309-3
Sequence 3, Application US/08460309
Patent No. 5837496
GENERAL INFORMATION:
APPLICANT: Engvall, Eva
APPLICANT: Leivo, Ilmo
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
FRAGMENTS AND USES THEREOF
```



NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,309  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/125,077  
FILING DATE: 22-SEP-1993  
APPLICATION NUMBER: US PCT/US 94/10730  
FILING DATE: 21-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/472,319  
FILING DATE: 30-JAN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/919,951  
FILING DATE: 27-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 9721  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6942 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-460-309-3

Query Match 76.0%; Score 15.2; DB 2; Length 6942;  
Best Local Similarity 85.0%; Fred. No. 96;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAGCCTGACAAAGCA 20  
DB 6163 GGACAAGCCTGACAAAGCA 6182

RESULT 6  
US-08-125-077-3  
Sequence 3, Application US/08125077  
Patent No. 5872231  
Patent No. 5872231 5840863  
GENERAL INFORMATION:  
APPLICANT: Engvall, Eva  
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
TITLE OF INVENTION: Fragments and Uses Thereof  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/125,077  
FILING DATE: 22-SEP-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US 94/10730  
FILING DATE: 21-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/472,319  
FILING DATE: 30-JAN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/919,951  
FILING DATE: 27-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 9721  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6942 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-125-077-3

Query Match 76.0%; Score 15.2; DB 2; Length 6942;  
Best Local Similarity 85.0%; Fred. No. 96;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAGCCTGACAAAGCA 20  
DB 6163 GGACAAGCCTGACAAAGCA 6182

RESULT 7  
US-09-252-991A-5635/C  
Sequence 5635, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 5635  
LENGTH: 1185  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-5635

Query Match 74.0%; Score 14.8; DB 4; Length 1185;  
Best Local Similarity 88.9%; Fred. No. 1.1e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAAGCCTGACAAAGC 18  
DB 67 GGACAAGCAGTGACAAAGC 50

RESULT 8  
US-09-489-847-113  
Sequence 113, Application US/09489847

Patent No. 6476195  
GENERAL INFORMATION:  
APPLICANT: Rosen et al  
TITLE OF INVENTION: 98 Human Secreted Proteins  
FILE REFERENCE: P2031P1  
CURRENT APPLICATION NUMBER: US/09/489,847  
CURRENT FILING DATE: 2000-01-24  
EARLIER APPLICATION NUMBER: PCT/US99/17130  
EARLIER FILING DATE: 1999-07-29  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30  
EARLIER APPLICATION NUMBER: 60/095,486  
EARLIER FILING DATE: 1998-08-05  
EARLIER APPLICATION NUMBER: 60/096,319  
EARLIER FILING DATE: 1998-08-12  
EARLIER APPLICATION NUMBER: 60/095,454  
EARLIER FILING DATE: 1998-08-06  
EARLIER APPLICATION NUMBER: 60/095,455  
NUMBER OF SEQ ID NOS: 376  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 113  
LENGTH: 2214  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (289)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-489-847-113

Query Match  
Best Local Similarity 74.0%; Score 14.8; DB 4; Length 2214;  
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCA 20  
DB 305 GGGRRAGGCTGACAGCA 324

RESULT 9  
US-09-252-991A-8300  
Sequence 8300, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 8300  
LENGTH: 360  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-8300

Query Match  
Best Local Similarity 72.0%; Score 14.4; DB 4; Length 360;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CAAGCCCTGACAGCC 19  
DB 245 CAAGCCCTGACAGCC 260

RESULT 10  
US-09-252-991A-8306

Sequence 8306, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 8306  
LENGTH: 417  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-8306

Query Match  
Best Local Similarity 72.0%; Score 14.4; DB 4; Length 417;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CAAGCCCTGACAGCC 19  
DB 302 CAAGCCCTGACAGCC 317

RESULT 11  
US-08-240-372-1  
Sequence 1, Application US/08240372  
Patent No. 5741665  
GENERAL INFORMATION:  
APPLICANT: KATO, ELIE K.  
TITLE OF INVENTION: LIGHT-REGULATED PROMOTERS FOR PRODUCTION  
OF HETEROLOGOUS PROTEINS IN FILAMENTOUS FUNGI  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESS: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W.  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/240,372  
FILING DATE: 10-MAY-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 3918-0003.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1237 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-240-372-1

Query Match  
Best Local Similarity 72.0%; Score 14.4; DB 1; Length 1237;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 AAGCCTGACAGCA 20  
Db 365 AAGCCTGACAGCA 380

## RESULT 12

US-09-564-805-221  
Sequence 221, Application US/09564805  
Patent No. 6333403  
GENERAL INFORMATION:  
APPLICANT: Tavitgian, Sean V.  
APPLICANT: Teng, David H.F.  
APPLICANT: Simard, Jacques M.  
APPLICANT: Rommens, Johanna M.  
APPLICANT: Myriad Genetics, Inc.  
TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes  
FILE REFERENCE: 2318-258  
CURRENT APPLICATION NUMBER: US/09/564,805  
CURRENT FILING DATE: 2000-05-05  
PRIOR APPLICATION NUMBER: US 60/107,468  
PRIOR FILING DATE: 1998-11-06  
PRIOR APPLICATION NUMBER: 09/434,382  
PRIOR FILING DATE: 1999-11-05  
NUMBER OF SEQ ID NOS: 240  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 221  
LENGTH: 2470  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(2466)  
US-09-564-805-221

Query Match 72.0%; Score 14.4; DB 4; Length 2470;  
Best Local Similarity 93.8%; Pred. No. 2.1e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 AAGCCTGACAGCA 20

Db 1479 AAGCCTGACAGTCA 1494

## RESULT 13

US-09-011-745-7  
Sequence 7, Application US/09011745  
Patent No. 6165715  
GENERAL INFORMATION:  
APPLICANT: Collins, Mary KL  
APPLICANT: Weiss, Robin A  
APPLICANT: Takeuchi, Yasuhiro  
APPLICANT: Cosset, Francois-Joic  
TITLE OF INVENTION: Depression systems  
FILE REFERENCE: 09/011,745  
CURRENT APPLICATION NUMBER: US/09/011,745  
CURRENT FILING DATE: 1998-06-22  
EARLIER APPLICATION NUMBER: PCT/GB96/02061  
EARLIER FILING DATE: 1996-08-23  
EARLIER APPLICATION NUMBER: GB9517263.1  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 7  
LENGTH: 6312  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Portion of  
OTHER INFORMATION: construct  
NAME/KEY: misc\_feature

LOCATION: (4058)  
OTHER INFORMATION: n is any nucleotide  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (4058)  
OTHER INFORMATION: n is any nucleotide  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (4060)  
OTHER INFORMATION: n is any nucleotide  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (4061)  
OTHER INFORMATION: n is any nucleotide  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (4246)  
OTHER INFORMATION: n is any nucleotide  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (4247)  
OTHER INFORMATION: n is any nucleotide  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (4248)  
OTHER INFORMATION: n is any nucleotide  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (4249)  
OTHER INFORMATION: n is any nucleotide  
US-09-011-745-7

Query Match 72.0%; Score 14.4; DB 3; Length 6312;  
Best Local Similarity 93.8%; Pred. No. 2.5e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 CAAGCCTGACAGCC 19

Db 3115 CAAGCCTGACAGCC 3130

RESULT 14  
US-09-315-127-4  
Sequence 4, Application US/09315127  
Patent No. 6448390  
GENERAL INFORMATION:  
APPLICANT: The University of Tennessee, c/o Richard Cox  
TITLE OF INVENTION: Stable Envelope Proteins for Retroviral, Viral and  
FILE REFERENCE: 44137-5023, U. of Tennessee  
CURRENT APPLICATION NUMBER: US/09/315,127  
CURRENT FILING DATE: 1999-05-20  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 8088  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: SEQ. ID NO. 3,  
OTHER INFORMATION: retroviral vector  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (5552)..(7552)  
US-09-315-127-4

Query Match 72.0%; Score 14.4; DB 4; Length 8088;  
Best Local Similarity 93.8%; Pred. No. 2.6e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 CAAGCCTGACAGCC 19

Db 7121 CAAGCCTGACAGCC 7136

## RESULT 15

US-08-716-351A-1

; Sequence 1, Application US/08716351A

; Patent No. 6033905

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Gibbon Ape Leukemia Virus-Based

; TITLE OF INVENTION: Retroviral Vectors

; NUMBER OF SEQUENCES: 5

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA: US/08/716,351A

; APPLICATION NUMBER: US/08/716,351A

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US94/03784

; FILING DATE: 06-APR-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Baselian, Kevin L.

; REGISTRATION NUMBER: 34,774

; REFERENCE/DOCKET NUMBER: 15280-128-1PC

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 543-9600

; TELEFAX: (415) 543-5043

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 8535 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 1..8535

; OTHER INFORMATION: /standard\_name= "GalV SEATO Genome"

US-08-716-351A-1

Query Match 72.0%; Score 14.4; DB 3; Length 8535;

Best Local Similarity 93.8%; Pred. No. 2.6e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CAAGCCTGACAGCC 19

Db 7491 CAAGCCTGACAGCC 7506

Search completed: August 1, 2003, 08:37:21  
Job time: 2.65066 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 02:21:37 ; Search time 1.65066 Seconds  
(without alignments)  
5347.959 Million cell updates/sec

Title: US-09-834-291-10

Sequence: 1 ggacaagccctgacaagcca 20

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfillseq1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15.2	76.0	1923	4 US-09-220-132-85	Sequence 85, Appl
2	15.2	76.0	3554	2 US-08-460-309-1	Sequence 1, Appl
3	15.2	76.0	3554	2 US-08-125-077-1	Sequence 1, Appl
4	15.2	76.0	3554	6 5444158-1	Patent No. 5444158
5	15.2	76.0	6942	6 US-08-460-309-3	Sequence 3, Appl
6	15.2	76.0	6942	2 US-08-125-077-3	Sequence 3, Appl
7	14.8	74.0	1185	4 US-09-252-991A-5635	Sequence 113, App
8	14.8	74.0	2214	4 US-09-489-847-113	Sequence 8306, Ap
9	14.4	72.0	360	4 US-09-252-991A-8300	Sequence 8306, Ap
10	14.4	72.0	417	4 US-09-252-991A-8306	Sequence 1, Appl
11	14.4	72.0	1237	1 US-08-240-372-1	Sequence 221, App
12	14.4	72.0	2470	3 US-09-564-805-221	Sequence 7, Appl
13	14.4	72.0	6312	3 US-09-011-745-7	Sequence 4, Appl
14	14.4	72.0	8088	4 US-09-315-127-4	Sequence 1, Appl
15	14.4	72.0	8535	3 US-08-716-351A-1	Sequence 120, App
16	14.4	72.0	36181	4 US-08-311-731A-120	Sequence 35, Appl
17	14.2	71.0	20	3 US-08-392-542-35	Sequence 35, Appl
18	14.2	71.0	20	3 US-08-894-337-35	Sequence 35, Appl
19	14.2	71.0	20	4 US-09-685-027-35	Sequence 18, Appl
20	14.2	71.0	30	1 US-08-347-792-18	Sequence 18, Appl
21	14.2	71.0	30	1 US-08-431-357-18	Sequence 18, Appl
22	14.2	71.0	30	2 US-08-697-221-29	Sequence 29, Appl
23	14.2	71.0	30	3 US-08-392-542-31	Sequence 31, Appl
24	14.2	71.0	30	3 US-08-894-337-31	Sequence 31, Appl
25	14.2	71.0	30	4 US-09-685-027-31	Sequence 18, Appl
26	14.2	71.0	30	5 PCT-US95-15353-18	Sequence 5, Appl
27	14.2	71.0	34	2 US-08-713-052-5	Sequence 5, Appl

28	14.2	71.0	72	3 US-09-276-531-93	Sequence 93, Appl
29	14.2	71.0	218	3 US-09-188-930-86	Sequence 86, Appl
30	14.2	71.0	218	4 US-09-312-283C-86	Sequence 86, Appl
31	14.2	71.0	538	3 US-08-979-586-3	Sequence 3, Appl
32	14.2	71.0	538	4 US-09-577-640-3	Sequence 28, Appl
33	14.2	71.0	780	1 US-08-445-909A-28	Sequence 11, Appl
34	14.2	71.0	1029	4 US-09-216-393B-11	Sequence 10642, A
35	14.2	71.0	1254	4 US-09-252-991A-10642	Sequence 68, Appl
36	14.2	71.0	1568	4 US-09-247-155-68	Sequence 10432, A
37	14.2	71.0	2169	4 US-09-252-991A-10432	Sequence 10524, A
38	14.2	71.0	2979	4 US-09-252-991A-10524	Sequence 16, Appl
39	14.2	71.0	3183	4 US-09-220-132-16	Sequence 571, App
40	14.2	71.0	3470	4 US-09-620-312D-571	Sequence 49, Appl
41	14.2	71.0	4057	3 US-08-894-997-49	Sequence 5, Appl
42	14.2	71.0	4533	3 US-08-726-214-5	Sequence 3, Appl
43	14.2	71.0	6942	2 US-08-460-309-3	Sequence 3, Appl
44	14.2	71.0	6942	2 US-08-125-077-3	Sequence 11, Appl
45	14.2	71.0	1230025	4 US-09-198-452A-1	Sequence 1, Appl

## ALIGNMENTS

RESULT 1  
US-09-220-132-85  
Sequence 85, Application US/09220132

Patent No. 6506607

GENERAL INFORMATION:

APPLICANT: Shyan, Andrew W.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER

FILE REFERENCE: 07334-074001

CURRENT APPLICATION NUMBER: US/09/220,132

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: US 60/079,303

PRIOR FILING DATE: 1998-03-25

PRIOR APPLICATION NUMBER: US 60/068,821

PRIOR FILING DATE: 1997-12-24

NUMBER OF SEQ ID NOS: 191

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 85

LENGTH: 1923

TYPE: DNA

ORGANISM: Homo sapiens

US-09-220-132-85

Query Match

Best Local Similarity 85.0%; Score 15.2; DB 4; Length 1923;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20

DB 1572 GCAAGCCCTGACAAAGCCA 1591

RESULT 2

US-08-460-309-1

Sequence 1, Application US/08460309

Patent No. 5837496

GENERAL INFORMATION:

APPLICANT: Engvall, Eva

TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin

FILE REFERENCE: 07334-074001

CURRENT APPLICATION NUMBER: US/09/220,132

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: US 60/079,303

PRIOR FILING DATE: 1998-03-25

PRIOR APPLICATION NUMBER: US 60/068,821

PRIOR FILING DATE: 1997-12-24

NUMBER OF SEQ ID NOS: 191

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 85

LENGTH: 1923

TYPE: DNA

ORGANISM: Homo sapiens

US-09-220-132-85

Query Match

Best Local Similarity 85.0%; Score 15.2; DB 4; Length 1923;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20

DB 1572 GCAAGCCCTGACAAAGCCA 1591

RESULT 2

US-08-460-309-1

Sequence 1, Application US/08460309

Patent No. 5837496

GENERAL INFORMATION:

APPLICANT: Engvall, Eva

TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin

FILE REFERENCE: 07334-074001

CURRENT APPLICATION NUMBER: US/09/220,132

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: US 60/079,303

PRIOR FILING DATE: 1998-03-25

PRIOR APPLICATION NUMBER: US 60/068,821

PRIOR FILING DATE: 1997-12-24

NUMBER OF SEQ ID NOS: 191

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 85

LENGTH: 1923

TYPE: DNA

ORGANISM: Homo sapiens

US-09-220-132-85

Query Match

Best Local Similarity 85.0%; Score 15.2; DB 4; Length 1923;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20

DB 1572 GCAAGCCCTGACAAAGCCA 1591

RESULT 2

US-08-460-309-1

Sequence 1, Application US/08460309

Patent No. 5837496

GENERAL INFORMATION:

APPLICANT: Engvall, Eva

TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin

FILE REFERENCE: 07334-074001

CURRENT APPLICATION NUMBER: US/09/220,132

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: US 60/079,303

PRIOR FILING DATE: 1998-03-25

PRIOR APPLICATION NUMBER: US 60/068,821

PRIOR FILING DATE: 1997-12-24

NUMBER OF SEQ ID NOS: 191

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 85

LENGTH: 1923

TYPE: DNA

ORGANISM: Homo sapiens

US-09-220-132-85

Query Match

Best Local Similarity 85.0%; Score 15.2; DB 4; Length 1923;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20

DB 1572 GCAAGCCCTGACAAAGCCA 1591

RESULT 2

US-08-460-309-1

Sequence 1, Application US/08460309

Patent No. 5837496

GENERAL INFORMATION:

APPLICANT: Engvall, Eva

TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin

FILE REFERENCE: 07334-074001

CURRENT APPLICATION NUMBER: US/09/220,132

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: US 60/079,303

PRIOR FILING DATE: 1998-03-25

PRIOR APPLICATION NUMBER: US 60/068,821

PRIOR FILING DATE: 1997-12-24

NUMBER OF SEQ ID NOS: 191

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 85

LENGTH: 1923

TYPE: DNA

ORGANISM: Homo sapiens

US-09-220-132-85

Query Match

Best Local Similarity 85.0%; Score 15.2; DB 4; Length 1923;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20

DB 1572 GCAAGCCCTGACAAAGCCA 1591

RESULT 2

US-08-460-309-1

Sequence 1, Application US/08460309

Patent No. 5837496

GENERAL INFORMATION:

APPLICANT: Engvall, Eva

TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin

FILE REFERENCE: 07334-074001

CURRENT APPLICATION NUMBER: US/09/220,132

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: US 60/079,303

PRIOR FILING DATE: 1998-03-25

PRIOR APPLICATION NUMBER: US 60/068,821

PRIOR FILING DATE: 1997-12-24

NUMBER OF SEQ ID NOS: 191

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 85

LENGTH: 1923

TYPE: DNA

ORGANISM: Homo sapiens

US-09-220-132-85

Query Match

Best Local Similarity 85.0%; Score 15.2; DB 4; Length 1923;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20

DB 1572 GCAAGCCCTGACAAAGCCA 1591

RESULT 2

US-08-460-309-1

Sequence 1, Application US/08460309

Patent No. 5837496

GENERAL INFORMATION:

APPLICANT: Engvall, Eva

TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin

FILE REFERENCE: 07334-074001

CURRENT APPLICATION NUMBER: US/09/220,132

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: US 60/079,303

PRIOR FILING DATE

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,309  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/125,077  
FILING DATE: 22-SEP-1993  
APPLICATION NUMBER: US PCT/US 94/10730  
FILING DATE: 21-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/472,319  
FILING DATE: 30-JAN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/919,951  
FILING DATE: 27-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 9721  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3554 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 10..3400  
US-08-460-309-1

Query Match 76.0%; Score 15.2; DB 2; Length 3554;  
Best Local Similarity 85.0%; Pred. No. 86;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAGCCTGACAGCA 20  
DB 183 GGACAAGCCTGACAGCA 202

RESULT 3  
US-08-125-077-1  
Sequence 1, Application US/08125077  
Patent No. 5872231  
Patent No. 5872231 5840863  
GENERAL INFORMATION:  
APPLICANT: Engvall, Eva  
APPLICANT: Leivo, Ilmo  
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/125,077  
FILING DATE: 22-SEP-1993  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US 94/10730  
FILING DATE: 21-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/472,319  
FILING DATE: 30-JAN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/919,951  
FILING DATE: 27-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 9721  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3554 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 10..3400  
US-08-125-077-1

Query Match 76.0%; Score 15.2; DB 2; Length 3554;  
Best Local Similarity 85.0%; Pred. No. 86;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAGCCTGACAGCA 20  
DB 183 GGACAAGCCTGACAGCA 202

RESULT 4  
544158-1  
Patent No. 5444158  
APPLICANT: ENGVAL, EVA; SANES, JOSHUA  
TITLE OF INVENTION: MEROSIN, NUCLEIC ACIDS ENCODING,  
FRAGMENTS AND USES THEREOF  
NUMBER OF SEQUENCES: 4  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/87,642  
FILING DATE: 08-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 587,689  
FILING DATE: 24-SEP-1990  
APPLICATION NUMBER: 472,319  
FILING DATE: 30-JAN-1990  
SEQ ID NO: 1:  
LENGTH: 3554  
544158-1

Query Match 76.0%; Score 15.2; DB 6; Length 3554;  
Best Local Similarity 85.0%; Pred. No. 86;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAGCCTGACAGCA 20  
DB 183 GGACAAGCCTGACAGCA 202

RESULT 5  
US-08-460-309-3  
Sequence 3, Application US/08460309  
Patent No. 5837496  
GENERAL INFORMATION:  
APPLICANT: Engvall, Eva  
APPLICANT: Leivo, Ilmo  
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
FRAGMENTS AND USES THEREOF

NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,309  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/125,077  
FILING DATE: 22-SEP-1993  
APPLICATION NUMBER: US PCT/US 94/10730  
FILING DATE: 21-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/472,319  
FILING DATE: 30-JAN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/919,951  
FILING DATE: 27-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 9721  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6942 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-460-309-3

Query Match 76.0%; Score 15.2; DB 2; Length 6942;  
Best Local Similarity 85.0%; Pred. No. 96;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20  
DB 6163 GGACAAAGCCGACAAAGCCA 6182

RESULT 6  
US-08-125-077-3  
Sequence 3, Application US/08125077  
Patent No. 5872231  
Patent No. 5872231 5840863  
GENERAL INFORMATION:  
APPLICANT: Engvall, Eva  
APPLICANT: leiyo, Ilmo  
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
TITLE OF INVENTION: Fragments and Uses Thereof  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/125,077  
FILING DATE: 22-SEP-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US 94/10730  
FILING DATE: 21-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/472,319  
FILING DATE: 30-JAN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/919,951  
FILING DATE: 27-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 9721  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6942 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-125-077-3

Query Match 76.0%; Score 15.2; DB 2; Length 6942;  
Best Local Similarity 85.0%; Pred. No. 96;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGCCA 20  
DB 6163 GGACAAAGCCGACAAAGCCA 6182

RESULT 7  
US-09-252-991A-5635/C  
Sequence 5635, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 5635  
LENGTH: 1185  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-5635

Query Match 74.0%; Score 14.8; DB 4; Length 1185;  
Best Local Similarity 88.9%; Pred. No. 1,1e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGACAAAGCCCTGACAAAGC 18  
DB 67 GGACAAAGCAGTGACAAAGC 50

RESULT 8  
US-09-489-847-113  
Sequence 113, Application US/09489847

Patent No. 6476195  
GENERAL INFORMATION:  
APPLICANT: Rosen et al  
TITLE OF INVENTION: 98 Human Secreted Proteins  
FILE REFERENCE: P2031P1  
CURRENT APPLICATION NUMBER: US/09/489,847  
CURRENT FILING DATE: 2000-01-24  
EARLIER APPLICATION NUMBER: PCT/US99/117130  
EARLIER FILING DATE: 1999-07-29  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30  
EARLIER APPLICATION NUMBER: 60/095,486  
EARLIER FILING DATE: 1998-08-05  
EARLIER APPLICATION NUMBER: 60/096,319  
EARLIER FILING DATE: 1998-08-12  
EARLIER APPLICATION NUMBER: 60/095,454  
EARLIER FILING DATE: 1998-08-06  
EARLIER APPLICATION NUMBER: 60/095,455  
EARLIER FILING DATE: 1998-08-06  
NUMBER OF SEQ ID NOS: 376  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 113  
LENGTH: 2214  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (289)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-489-847-113

Query Match 74.0%; Score 14.8; DB 4; Length 2214;  
Best Local Similarity 80.0%; Pred. No. 1.3e+02;  
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGACAGCCCTGACAGCCA 20  
||:|||||  
Db 305 GGGRAGCCCTGACAGCCA 324

RESULT 9  
US-09-252-991A-8300  
Sequence 8300, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 8300  
LENGTH: 360  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-8300

Query Match 72.0%; Score 14.4; DB 4; Length 360;  
Best Local Similarity 93.8%; Pred. No. 1.5e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CAAGCCCTGACAGCC 19  
|||||  
Db 245 CAAGCCCTGACAGCC 260

RESULT 10  
US-09-252-991A-8306

Sequence 8306, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 8306  
LENGTH: 417  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-8306

Query Match 72.0%; Score 14.4; DB 4; Length 417;  
Best Local Similarity 93.8%; Pred. No. 1.5e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CAAGCCCTGACAGCC 19  
|||||  
Db 302 CAAGCCCTGACAGCC 317

RESULT 11  
US-08-240-372-1  
Sequence 1, Application US/08240372  
Patent No. 5741665  
GENERAL INFORMATION:  
APPLICANT: KATO, ELIE K.  
APPLICANT: STUART, W. DORSEY  
TITLE OF INVENTION: LIGHT-REGULATED PROMOTERS FOR PRODUCTION  
TITLE OF INVENTION: OF HETEROLOGOUS PROTEINS IN FILAMENTOUS FUNGI  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W.  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/240,372  
FILING DATE: 10-MAY-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 3918-0003.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1237 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-240-372-1

Query Match 72.0%; Score 14.4; DB 1; Length 1237;  
Best Local Similarity 93.8%; Pred. No. 1.9e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;



QY 5 AAGCCTGACAGCCA 20  
|||||  
Db 365 AAGCCTGACAGACA 380

## RESULT 12

US-09-564-805-221  
Sequence 221, Application US/09564805  
Patent No. 6333403

## GENERAL INFORMATION:

APPLICANT: Tavtigian, Sean V.

APPLICANT: Teng, David H.F.

APPLICANT: Simard, Jacques

APPLICANT: Rommens, Johanna M.

APPLICANT: Myriad Genetics, Inc.

TITLE OF INVENTION: Chromosome 17p-linked Prostate Cancer Susceptibility

TITLE OF INVENTION: Gene and a Paralog and Orthologous Genes

FILE REFERENCE: 2318-258

CURRENT APPLICATION NUMBER: US/09/564,805

PRIOR FILING DATE: 2000-05-05

PRIOR APPLICATION NUMBER: US 60/107,468

PRIOR FILING DATE: 1998-11-06

PRIOR APPLICATION NUMBER: 09/434,382

PRIOR FILING DATE: 1999-11-05

NUMBER OF SEQ ID NOS: 240

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO: 221

LENGTH: 2470

TYPE: DNA

ORGANISM: Mus musculus

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(2466)

US-09-564-805-221

Query Match 72.0%; Score 14.4; DB 4; Length 2470;

Best Local Similarity 93.8%; Pred. No. 2.1e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AAGCCTGACAGCCA 20  
|||||  
Db 1479 AAGCCTGACAGTCA 1494

## RESULT 13

US-09-011-745-7

Sequence 7, Application US/09011745

Patent No. 6165715

GENERAL INFORMATION:

APPLICANT: Collins, Mary KL

APPLICANT: Weise, Robin A

APPLICANT: Takeuchi, Yasuhiro

APPLICANT: Cosset, Francois-Loic

TITLE OF INVENTION: Expression systems

FILE REFERENCE: 09/011,745

CURRENT APPLICATION NUMBER: US/09/011,745

PRIOR FILING DATE: 1998-06-22

EARLIER APPLICATION NUMBER: PCT/GB96/02061

EARLIER FILING DATE: 1996-08-23

EARLIER APPLICATION NUMBER: GB9517263.1

NUMBER OF SEQ ID NOS: 29

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO: 7

LENGTH: 6312

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE: OTHER INFORMATION: Description of Artificial Sequence: Portion of

OTHER INFORMATION: construct

NAME/KEY: misc\_feature

LOCATION: (4058)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (4059)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (4060)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (4061)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (4062)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (4063)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (4064)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (4065)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (4066)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (4067)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (4068)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (4069)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (4070)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (4071)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (4072)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (4073)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (4074)

OTHER INFORMATION: n is any nucleotide

FEATURE:

NAME/KEY: misc\_feature

QY 4 CAAGCCTGACAGCC 19  
|||||  
Db 3115 CAAGCCTGACAGCC 3130

Query Match 72.0%; Score 14.4; DB 3; Length 6312;  
Best Local Similarity 93.8%; Pred. No. 2.5e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 14  
US-09-315-127-4  
Sequence 4, Application US/09315127  
Patent No. 6448390

GENERAL INFORMATION:

APPLICANT: The University of Tennessee, c/o Richard Cox

TITLE OF INVENTION: Stable Envelope Proteins for Retroviral, Viral and

TITLE OF INVENTION: Liposome Vectors and Use in Gene and Drug Therapy

FILE REFERENCE: 44137-5023, U. of Tennessee

CURRENT APPLICATION NUMBER: US/09/315,127

PRIOR FILING DATE: 1999-05-20

NUMBER OF SEQ ID NOS: 23

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO: 4

LENGTH: 8088

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE: OTHER INFORMATION: Description of Artificial Sequence: SEQ. ID NO. 3,

OTHER INFORMATION: retroviral vector

FEATURE:

NAME/KEY: CDS

LOCATION: (5552)..(7552)

US-09-315-127-4

Query Match 72.0%; Score 14.4; DB 4; Length 8088;

Best Local Similarity 93.8%; Pred. No. 2.6e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CAAGCCTGACAGCC 19  
|||||  
Db 7121 CAAGCCTGACAGCC 7136

## RESULT 15

US-08-716-351A-1

; Sequence 1, Application US/08716351A

; Patent No. 6033905

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Gibbon Ape Leukemia Virus-Based

; TITLE OF INVENTION: Retroviral Vectors

; NUMBER OF SEQUENCES: 5

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/716.351A

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US94/03784

; FILING DATE: 06-APR-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Baselian, Kevin L.

; REGISTRATION NUMBER: 34,774

; REFERENCE/DOCKET NUMBER: 15280-128-1PC

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 543-9600

; TELEFAX: (415) 543-5043

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 8535 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 1..8535

; OTHER INFORMATION: /standard\_name="GalV SEATO Genome"

; US-08-716-351A-1

Query Match 72.0%; Score 14.4; DB 3; Length 8535;

Best local similarity 93.8%; Pred. No. 2.6e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CAAGCCTGACACGCC 19

DB 7491 CAAGCCTGACACGCC 7506

Search completed: August 1, 2003, 08:37:09  
Job time: 2.65066 secs